

① RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

⑪ N° de publication : 2 692 592
(à n'utiliser que pour les
commandes de reproduction)

⑫ N° d'enregistrement national : 92 07493

⑬ Int Cl⁸ : C 12 N 15/31, 1/21, C 12 P 21/02(C 12 N 15/31, C 12 R 1:36)

⑭ DEMANDE DE BREVET D'INVENTION

A1

⑮ Date de dépôt : 19.06.92.

⑯ Priorité :

⑰ Date de la mise à disposition du public de la demande : 24.12.93 Bulletin 93/51.

⑱ Liste des documents cités dans le rapport de recherche préliminaire : *Se reporter à la fin du présent fascicule.*

⑲ Références à d'autres documents nationaux apparentés :

⑳ Demandeur(s) : PASTEUR MERIEUX Sérums et Vaccins société anonyme — FR et TRANSGENE (S.A.) société anonyme — FR.

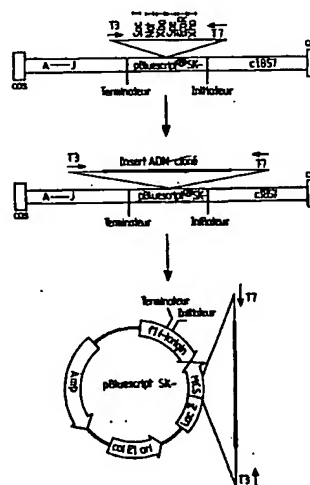
㉑ Inventeur(s) : Jacobs Eric, Legrain Michèle, Mazarin Véronique, Bouchon-Theisen Bernadette, Shryvers Anthony B. et Bloch Marie-Aline.

㉒ Titulaire(s) :

㉓ Mandataire : Cabinet Lemoine et Bernasconi.

㉔ Fragments d'ADN codant pour les sous-unités du récepteur de la transferrine de *Neisseria meningitidis* et procédés les exprimant.

㉕ La présente invention a pour objet un fragment d'ADN codant pour une protéine capable d'être reconnue par un antisérum anti-récepteur de la transferrine de la souche de *N. meningitidis* IM2394 ou IM2169 ainsi d'un procédé d'obtention de ladite protéine par voie recombinante. A titre d'exemple, un tel fragment d'ADN code pour la sous-unité *tbp1* de la souche IM2394 ou IM2169 ou pour la sous-unité *tbp2* de la souche IM2394 ou IM2169.



FR 2 692 592 - A1



BEST AVAILABLE COPY

La présente invention a pour objet des fragments d'ADN de *Neisseria meningitidis* codant pour les sous-unités du récepteur de la transferrine ainsi qu'un procédé de fabrication de chacune des sous-unités par voie recombinante.

5

D'une manière générale, les méningites sont soit d'origine virale, soit d'origine bactérienne. Les bactéries principalement responsables sont : *N. meningitidis* et *Haemophilus influenzae*, respectivement impliquées dans environ 40 et 50 % des cas de méningites bactériennes.

10

On dénombre en France, environ 600 à 800 cas par an de méningites à *N. meningitidis*. Aux Etats-Unis, le nombre de cas s'élève à environ 2 500 à 3 000 par an.

15

L'espèce *N. meningitidis* est subdivisée en sérogroupes selon la nature des polysaccharides capsulaires. Bien qu'il existe une douzaine de sérogroupes, 90 % des cas de méningites sont attribuables à 3 sérogroupes : A, B et C.

20

Il existe des vaccins efficaces à base de polysaccharides capsulaires pour prévenir les méningites à *N. meningitidis* sérogroupes A et C. Ces polysaccharides tels quels ne sont que peu ou pas immunogéniques chez les enfants de moins de 2 ans et n'induisent pas de mémoire immunitaire. Toutefois, ces inconvénients peuvent être surmontés en conjuguant ces polysaccharides à une protéine porteuse.

25

Par contre, le polysaccharide de *N. meningitidis* groupe B n'est pas ou peu immunogène chez l'homme, qu'il soit sous forme conjuguée ou non. Ainsi, il apparaît hautement souhaitable de rechercher un vaccin à l'encontre des méningites induites par *N. meningitidis* notamment du séro groupe B autre qu'un vaccin à base de polysaccharide.

30

A cette fin, différentes protéines de la membrane externe de *N. meningitidis* ont déjà été proposées. Il s'agit en particulier du récepteur membranaire de la transferrine humaine.

35

D'une manière générale, la grande majorité des bactéries ont besoin de fer pour leur croissance et elles ont développé des systèmes spécifiques d'acquisition de ce métal. En ce qui concerne notamment *N. meningitidis* qui est un pathogène strict de l'homme, le fer ne peut être prélevé qu'à partir de protéines humaines de transport du fer telles que la transferrine et la lactoferrine puisque la quantité de fer sous forme libre est négligeable chez l'homme (de l'ordre de 10^{-18} M), en tout cas insuffisante pour permettre la croissance bactérienne.

Ainsi, *N. meningitidis* possède un récepteur de la transferrine humaine et un récepteur de la lactoferrine humaine qui lui permettent de fixer ces protéines chélatrices du fer et de capter par la suite le fer nécessaire à sa croissance.

Le récepteur de la transferrine de la souche *N. meningitidis* B16B6 a été purifié par Schryvers et al (WO 90/12591) à partir d'un extrait membranaire. Cette protéine telle que purifiée apparaît essentiellement constituée de 2 types de polypeptides : un polypeptide d'un poids moléculaire apparent élevé de 100 kD et un polypeptide d'un poids moléculaire apparent moindre d'environ 70 kD, telles que révélés après électrophorèse sur gel de de polyacrylamide en présence de SDS.

Le produit de la purification notamment mise en oeuvre par Schryvers est par définition arbitraire et pour les besoins de la présente demande de brevet, appelé récepteur de la transferrine et les polypeptides le constituant, des sous-unités. Dans la suite du texte, les sous-unités de poids moléculaire élevé et de poids moléculaire moindre sont respectivement appelées Tbp1 et Tbp2.

Toutefois, le procédé de purification décrit par Schryvers et al ne peut pas être utilisé pour la production à grande échelle du récepteur de la transferrine. La préparation industrielle de ce récepteur sous forme purifiée passe nécessairement par une étape de production à l'aide d'un système d'expression hétérologue.

A cette fin, l'invention se propose de fournir les fragments d'ADN codant pour les sous-unités du récepteur de la transferrine de *N. meningitidis*.

5 D'autre part, depuis les travaux pionniers de Schryvers et al, on a découvert qu'il existait en fait au moins 2 types de souches qui diffèrent par la constitution de leurs récepteurs de la transferrine respectifs. Ceci a été mis en évidence en étudiant des extraits membranaires de plusieurs dizaines de souches de *N. meningitidis* d'origines variées. Ces extraits membranaires ont tout d'abord été soumis à une électrophorèse sur gel de polyacrylamide en
10 présence de SDS, puis électrotransférés sur feuilles de nitrocellulose. Ces feuilles de nitrocellulose ont été incubées :

- 15 a) en présence d'un antisérum de lapin dirigé contre le récepteur de la transferrine purifié à partir de la souche *N. meningitidis* B16B6, aussi appelée IM2394 ;
- b) en présence d'un antisérum de lapin dirigé contre le récepteur de la transferrine purifié à partir de la souche *N. meningitidis* IM2169 ; ou
- 20 c) en présence de la transferrine humaine conjuguée à la peroxydase.

En ce qui concerne a) et b), la reconnaissance des sous-unités du récepteur de la transferrine est révélée par addition d'un anticorps anti-immunoglobulines de lapin couplé à la peroxydase, puis par addition du
25 substrat de cette enzyme.

Les tableaux I et II ci-dessous indiquent le profil de certaines souches représentatives tel qu'il apparait sur gel de polyacrylamide à 7,5 % après électrophorèse en présence de SDS ; les bandes sont caractérisées par leur
30 poids moléculaires apparents exprimés en kilodaltons (kD) :

	Souches			
Tableau I	2394 (B; 2a;P1.2:L2,3) 2228 (B; nd) 2170 (B; 2a;P1.2:L3)	2234 (Y; nd) 2154 (C; nd) 2448 (B; nd)	550 (C; 2a;) 179 (C; 2a;P1.2)	
Détection avec l'antisérum anti-récepteur 2394	93 68	93 69	99 69	
Détection avec l'antisérum anti-récepteur 2169	93	93	99	
Détection avec la transferrine peroxydase	68	69	69	

N.B. : Entre parenthèse sont indiqués dans l'ordre le sérotype, le sous-type et l'immunotype.

	Souches								
Tableau II	2169 (B:9:P1.9)	1000 (B:nd)	1604 (B:nd)	132 (C:15:P1.16)	1001 (A:4:P1.9)	876 (B:19:P1.6)	1951 (A:nd)	2449 (B:nd)	867 (B:2b:P1.2)
Détection avec l'antisérum anti-récepteur 2394	96	98	98	98	98	96	94	94	93
Détection avec l'antisérum anti-récepteur 2169	96 87	98 85	98 83	98 81	98 79	96 88	94 87	94 85	93 85
Détection avec la transferrine- peroxydase	87	85	83	81	79	88	87	85	85

N.B. : Entre parenthèse sont indiqués dans l'ordre le sérotype, le sous-type et l'immunotype.

Les résultats répertoriés dans les 2 premières lignes des tableaux montrent qu'il existe 2 types de souches :

5 Le premier type (Tableau I) correspond à des souches qui possèdent un récepteur dont les 2 sous-unités dans les conditions expérimentales utilisées, sont reconnues par l'antisérum anti-récepteur IM2394 tandis que seule la sous-unité de haut poids moléculaire est reconnue par l'antisérum anti-récepteur IM2169.

10 Le second type (Tableau II) correspond à des souches qui possèdent un récepteur dont les 2 sous-unités dans les conditions expérimentales utilisées, sont reconnues par l'antisérum anti-récepteur IM2169 tandis que seule la sous-unité de haut poids moléculaire est reconnue par l'antisérum anti-récepteur IM2394.

15 En conséquence, il existe une diversité antigénique au niveau de la sous-unité de moindre poids moléculaire. Cette diversité est toutefois restreinte puisqu'elle se résout en 2 grands types, contrairement à ce qui est suggéré par Griffiths et al, FEMS Microbiol. Lett. (1990) 69 : 31.

20 En vertu de ces constatations, on pouvait supposer qu'un vaccin efficace à l'encontre de toutes les infections à *N. meningitidis* pourrait être constitué de manière suffisante, de la sous-unité de haut poids moléculaire, quelle que soit la souche d'origine du récepteur, puisque cette dernière est reconnue par
25 les 2 types d'antisérums. Toutefois, il semble que cela ne puisse être le cas dans la mesure où la sous-unité de haut poids moléculaire ne serait pas capable d'induire la production d'anticorps de type neutralisant. Seule la plus petite des 2 sous-unités du récepteur serait capable de remplir cette fonction. Puisque cette sous-unité de moindre poids moléculaire se caractérise par une
30 variation antigénique significative du premier type au deuxième type de souche, un seul type de récepteur de la transferrine ne devrait pas être suffisant pour vacciner contre toutes les infections à *N. meningitidis*. Par conséquent, un vaccin devra contenir au moins la sous-unité de moindre poids moléculaire de chacune des souches IM2394 et IM2169 ou de leurs
35 équivalents respectifs et, de manière optionnelle, la sous-unité de haut poids moléculaire d'au moins une souche de *N. meningitidis*.

C'est pourquoi l'invention fournit un fragment d'ADN isolé codant pour une protéine capable d'être reconnue par un antisérum anti-récepteur de la souche de *N. meningitidis* IM2394 ou IM2169.

5

Un tel fragment d'ADN peut notamment comprendre une séquence nucléotidique codant pour une séquence d'acides aminés homologue à celle telle que montrée :

- 10 - dans l'identificateur de séquence (SEQ ID NO : 1) n° 1, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579 ;
- dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en
15 position 1 et finissant avec le résidu phénylalanine en position 884 ;
- dans le SEQ ID NO : 3, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887 ; ou
- 20 - dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691.

A titre indicatif, on précise qu'un fragment d'ADN selon l'invention peut en outre comprendre une séquence nucléotidique additionnelle codant
25 pour n'importe quelle autre séquence d'acides aminés ; les deux séquences nucléotidiques considérées, formant un cadre ouvert de lecture de manière à coder pour une protéine hybride ou un précurseur.

De manière avantageuse, un fragment d'ADN selon l'invention peut
30 être sélectionné parmi :

- i) Un premier fragment d'ADN isolé, ayant une séquence nucléotidique codant pour une protéine ayant une séquence d'acides aminés homologue à celle telle que montrée dans le SEQ ID NO : 1, commençant avec le
35 résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579.

- 5 ii) Un deuxième fragment d'ADN isolé ayant une séquence nucléotidique codant pour une protéine ayant une séquence d'acides aminés à celle telle que montrée dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 884.
- 10 iii) Un troisième fragment d'ADN isolé ayant une séquence nucléotidique codant pour une protéine ayant une séquence d'acides aminés à celle telle que montrée dans le SEQ ID NO : 3, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887.
- 15 iv) Un quatrième fragment d'ADN isolé ayant une séquence nucléotidique codant pour une protéine ayant une séquence d'acides aminés à celle telle que montrée dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691.

20 Par "séquence d'acides aminés homologue", on entend une séquence présentant un degré d'homologie d'au moins 75 %, de manière avantageuse d'au moins 80 %, de manière préférée d'au moins 90 %, de manière tout à fait préférée de 100 %, avec la séquence d'acides aminés que l'on cite en référence. On notera que le terme "homologue" tel que défini inclut le cas particulier de l'identité.

25

30 Le degré d'homologie peut être aisément calculé en alignant les séquences de manière à obtenir le degré maximal d'homologie ; pour ce faire, il peut être nécessaire d'introduire artificiellement des emplacements vacants, comme cela est illustré dans la figure 7. Une fois que l'alignement optimal est réalisé, le degré d'homologie est établi en comptabilisant toutes les positions dans lesquelles les acides aminés des deux séquences se retrouvent à l'identique, par rapport au nombre total de positions.

Il serait fastidieux de décrire des séquences homologues autrement que de manière générique, en raison du trop grand nombre de combinaisons. L'homme du métier connaît toutefois les règles générales qui permettent de remplacer un acide aminé par un autre sans abolir la fonction biologique ou immunologique d'une protéine.

Un fragment d'ADN isolé et tout à fait préféré a une séquence nucléotidique codant pour :

- 10 i) La sous-unité Tbp1 de la souche IM2394 dont la séquence en acides aminés est telle que montrée dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 884 ;
- 15 ii) La sous-unité Tbp2 de la souche IM2394 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 1, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579 ;
- 20 iii) La sous-unité Tbp1 de la souche IM2169 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 3, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887 ; ou
- 25 iv) La sous-unité Tbp2 de la souche IM2169 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691 .

30 Le récepteur de la transferrine étant une protéine membranaire, chacune de ses sous-unités est initialement produite sous forme d'un précurseur constitué d'un peptide signal associé en position N-terminale, à la forme mature.

C'est pourquoi l'invention a aussi pour objet un bloc d'ADN isolé codant pour un peptide signal dont la séquence d'acides aminés présente un degré d'homologie d'au moins 80 %, de manière préférée de 100 %, avec la séquence montrée dans :

5

i) le SEQ ID NO : 2, commençant avec le résidu méthionine en position - 24 et finissant avec le résidu alanine en position - 1 ;

10

ii) le SEQ ID NO : 3, commençant avec le résidu méthionine en position - 24 et finissant avec le résidu alanine en position - 1 ; ou

iii) le SEQ ID NO : 4, commençant avec le résidu méthionine en position - 20 et finissant avec le résidu alanine en position - 1.

15

Un fragment d'ADN selon l'invention peut être aussi sélectionné parmi un cinquième, sixième, septième et huitième fragments d'ADN codant respectivement pour un précurseur dont la séquence d'acides aminés est homologue à la séquence présentée dans le SEQ ID NO : 1, 2, 3 ou 4.

20

Par "fragment ou bloc d'ADN isolé" on entend un fragment ou bloc d'ADN d'origine génomique qui est i) inséré dans un vecteur viral ou plasmidique ou ii) placé sous le contrôle d'un promoteur qui lui est hétérologue.

25

De plus, le bloc d'ADN codant pour le peptide signal selon l'invention est, en outre, considéré comme isolé lorsque ce bloc d'ADN est associé à un fragment d'ADN codant pour une protéine hétérologue au peptide signal ; de manière à former un cadre de lecture ouvert codant pour un précurseur hybride.

30

L'invention concerne aussi une cassette d'expression qui comprend au moins un fragment d'ADN selon l'invention, placé sous le contrôle d'éléments capables d'assurer son expression dans une cellule-hôte appropriée.

Dans la cassette d'expression, le premier, deuxième, troisième ou quatrième fragment d'ADN selon l'invention peut être ou non associé à un bloc d'ADN codant pour un peptide signal hétérologue à la protéine codée par ledit fragment d'ADN, selon que l'on recherche ou non la sécrétion de la protéine. De préférence, cette sécrétion sera recherchée.

Les éléments indispensables à l'expression d'un fragment d'ADN selon l'invention sont un promoteur de transcription, des codons de début et de fin de traduction et, de manière optionnelle, un terminateur de transcription. Le promoteur peut être constitutif ou inductible. On indique que le fragment d'ADN codant pour la sous-unité Tbp2 de la souche IM2394 semble être toxique pour une cellule hétérologue, notamment pour *E. coli*. Dans ce cas là, il pourrait être préférable d'utiliser un promoteur inductible.

Des éléments tels qu'un bloc d'ADN codant pour un peptide signal hétérologue (région signal) ou un promoteur existent déjà en assez grand nombre et sont connus de l'homme du métier. Ses compétences générales lui permettront de choisir une région signal ou un promoteur particulier qui seront adaptés à la cellule-hôte dans laquelle il envisage l'expression.

Enfin, l'invention fournit un procédé de fabrication d'un peptide, d'un polypeptide ou d'une protéine capables d'être reconnus par un antisérum anti-récepteur de la souche de *N. meningitidis* IM2394 ou IM2169 qui comprend l'acte de cultiver une cellule-hôte comportant une cassette d'expression selon l'invention ; ainsi que le peptide, le polypeptide ou la protéine produit par ce procédé et les compositions vaccinales les contenant.

Aux fins du procédé selon l'invention, la cellule-hôte peut être une cellule de mammifère, une bactérie ou une levure ; ces deux dernières étant préférées. Là aussi, le choix d'une lignée particulière est à la portée de l'homme du métier.

Afin de déterminer l'objet de la présente invention, on précise que les souches de *N. meningitidis* IM2394 et IM2169 sont publiquement disponibles auprès de la Collection Nationale de Culture des Microorganismes (CNCM), Institut Pasteur, 25 rue du Dr Roux 75015 Paris sous les numéros
5 d'enregistrement respectifs LNP N 1511 et LNP N 1520.

Un antisérum spécifique du récepteur de la transferrine de la souche de *N. meningitidis* IM2394 ou IM2169 peut être obtenu tel que décrit dans les exemples ci-après.
10

L'invention est décrite plus en détails dans les exemples ci-après et par référence aux Figures 1 à 8.

La Figure 1 représente la structure du phage lambda ZAP II et schématise la méthodologie de clonage y afférent. Lambda ZAP II est un vecteur d'insertion équipé de sites de clonage multiples localisés dans la partie plasmidique (pBluescript SK). Cette partie plasmidique peut être excisée *in vivo* par co-infection avec un phage-helper et convertie en vecteur plasmidique. Si une séquence codante est fusionnée en phase à lacZ ou si un
15 fragment d'ADN cloné comporte un promoteur fonctionnel dans *E. coli*, il peut y avoir production d'une protéine d'intérêt qui pourra être détectée à l'aide d'anticorps spécifiques.
20

La Figure 2 présente la structure du plasmide pTG1265. pTG1265 dérive du plasmide pGB2 (Churchward et al, Gene (1984) 31 : 165) comme
25 suit : pGB2 est digéré par *EcoRI* et *HindIII*, traité à la polymérase Klenow puis ligué au fragment *SspI* - *PvuII* de 1 kb issu de pT7T3 184 (Mead et al, Protein Engineering (1986) 1 : 67 ; Pharmacia) qui comporte fl-ori, la séquence lacZ, les promoteurs T3 et T7 ainsi que des sites multiples de
30 clonage.

La Figure 3 présente la carte génomique de la région d'ADN de la souche IM2394 comportant les séquences codant pour Tbp1 et Tbp2 ainsi que les différents fragments qui ont été clonés. B = *BamHI* ; E = *EcoRI* ; H
35 = *HincII* ; R = *EcoRV* ; X = *XbaI* ; C = *ClaI*.

La Figure 4 présente la carte génomique de la région d'ADN de la souche IM2169 comportant les séquences codant pour Tbp1 et Tbp2 ainsi que les différents fragments qui ont été clonés. C = *ClaI* ; H = *HincII* ; M = *MluI* ; X = *XbaI* ; ? = position imprécise.

5

La Figure 5 présente la structure du plasmide para13. para13 est un plasmide capable de se répliquer dans *E. coli* qui comporte le promoteur de l'opéron arabinose BAD (ParaB) de *Salmonella typhimurium* (modifié au niveau de la TATA box), ainsi que le gène AraC. En aval du promoteur ParaB se trouve des sites multiples d'insertion. La série des plasmides para est décrite par Cagnon et al, Prot. Eng. (1991) 4 : 843.

10

15

La Figure 6 représente la méthodologie qui a été mise en oeuvre pour construire le vecteur d'expression pTG3786.

La Figure 7 compare les séquences d'acides aminés prédites des sous-unités Tbp1 des souches IM2394 et IM2169. Le degré d'homologie peut être estimé à environ 76 %.

20

La Figure 8 compare les séquences d'acides aminés prédites des sous-unités Tbp2 des souches IM2394 et IM2169. Le degré d'homologie peut être estimé à environ 47 %.

25

La Figure 9 représente la méthodologie qui a été mise en oeuvre pour construire le vecteur d'expression pTG3779.

EXEMPLE 1 : Clonage des fragments d'ADN codant pour les sous-unités Tbp1 et Tbp2 du récepteur de la transferrine de la souche IM2394

5 **1A - Culture de la souche et purification du récepteur de la transferrine**

Un lyophilisat de la souche *N. meningitidis* IM2394 est repris dans environ 1 ml de bouillon Muller-Hinton (BMH, Difco). La suspension bactérienne est ensuite étalée sur le milieu solide Muller-Hinton contenant
10 du sang cuit (5 %).

Après 24 hr d'incubation à 37°C dans une atmosphère contenant 10 % de CO₂, la nappe bactérienne est recueillie pour ensemer 150 ml de BMH pH 7,2, répartis en 3 erlens de 250 ml. L'incubation est poursuivie
15 pendant 3 hr à 37°C sous agitation. Chacune des 3 cultures ainsi réalisées permet d'ensemencer 400 ml de BMH pH 7,2 supplémenté avec 30 µm de Ethylènediamine - Di (O-Hydroxyphenyl - acetic acid (EDDA, Sigma), qui est un agent chélatant du fer sous forme libre.

20 Après 16 hr de culture à 37°C sous agitation, les cultures sont contrôlées pour leur pureté par observation au microscope après une coloration de Gram. La suspension est centrifugée, le culot contenant les germes est pesé et conservé à -20°C.

25 La purification est mise en oeuvre essentiellement selon la méthode décrite par Schryvers et al (supra), comme suit :

Le culot bactérien est décongelé, puis remis en suspension dans 200 ml de tampon Tris HCl 50 mM, pH 8.0 (tampon A). La suspension est
30 centrifugée pendant 20 min à 15 000 xg à 4°C. Le culot est récupéré, puis remis en suspension dans du tampon A à la concentration finale de 150 g/l. Des fractions de 150 ml sont traitées pendant 8 min à 800 bars dans un lyseur de cellules travaillant sous haute pression (Rannie, modèle 8.30H). Le lysat cellulaire ainsi obtenu est centrifugé pendant 15 min à 4°C à 15 000 xg. Le
35 surnageant est récupéré, puis centrifugé pendant 75 min à 4°C à 200 000 xg. Après élimination du surnageant, le culot est repris dans du tampon A et

après dosage de protéines selon Lowry, la concentration de la suspension est ajustée à 5 mg/ml.

5 A 1,4 ml de la suspension de membranes on ajoute 1,75 mg de transferrine humaine biotinylée selon le procédé décrit par Schryvers. La concentration finale de la fraction membranaire est de 4 mg/ml. Le mélange est incubé 1 heure à 37°C puis centrifugé à 100 000 xg pendant 75 minutes à 4°C. Le culot de membranes est repris par le tampon A contenant du NaCl 0,1M et incubé pendant 60 minutes à température ambiante.

10

Après solubilisation, on ajoute à cette suspension un certain volume de N-Lauroyl Sarkosine à 30 % (p/v) et d'EDTA 500 mM de façon que les concentrations finales en Sarkosyl et EDTA soient de 0,5 % et 5 mM respectivement. Après une incubation de 15 minutes à 37°C sous agitation,
15 on ajoute 1 ml de résine streptavidine-agarose (Pierce) préalablement lavée en tampon A. La suspension est incubée 15 minutes à température ambiante puis centrifugée à 1 000 xg pendant 10 minutes. La résine est ensuite conditionnée dans une colonne et l'éluat direct est éliminé.

20

La résine est lavée par 3 volumes de colonne de tampon Tris-HCl 50 mM pH 8.0 contenant NaCl 1M, EDTA 10 mM Sarkosyl 0,5 % (tampon B) puis par un volume de colonne de tampon B contenant de la guanidine-HCl 750 mM. Le récepteur de la transferrine est ensuite élué par le tampon B contenant de la guanidine-HCl 2M. L'éluat est collecté en fractions, dans des
25 tubes contenant un volume identique de Tris HCl 50 mM, pH 8.0, NaCl 1M. La densité optique à 280 nm de l'éluat est mesurée en sortie de colonne à l'aide d'un détecteur UV.

30

Les fractions correspondant au pic d'élution sont recueillies, dialysées contre du tampon phosphate 10 mM, pH 8,0 contenant du Sarkosyl 0,05 % et lyophilisées. Le lyophilisat est repris dans de l'eau à une concentration 10 fois supérieure. La solution est dialysée une seconde fois contre du tampon phosphate 50 mM pH 8,0 contenant du Sarkosyl 0,05 % (tampon C) puis la solution est filtrée sur une membrane de porosité 0,22 µm.

35

Le contenu en protéines est déterminé et ajusté à 1 mg/ml par

addition de tampon C, sous conditions aseptiques. Cette préparation est conservée à -70°C.

1B - Préparation d'un antisérum spécifique du récepteur de la transferrine

5 Des lapins néo-zélandais albinos reçoivent par voie sous-cutanée et intramusculaire 100 µg du récepteur IM2394 en présence d'adjuvant complet de Freund. 21 jours et 42 jours après la première injection, les lapins reçoivent à nouveau 100 µg du récepteur purifié mais ces fois-ci en présence
10 d'adjuvant incomplet de Freund. 15 jours après la dernière injection, le sérum des animaux est prélevé, puis décomplémenté et filtré sur une membrane de porosité 0,45 µm. Le filtrat est par la suite épuisé par contact avec la souche IM2394 qui pour se faire, a été cultivée au préalable en présence de fer sous forme libre (dans ces conditions, la synthèse du
15 récepteur de la transferrine est réprimée). Les modalités de contact sont comme suit : 10 ml du filtrat sont ajoutés à 10¹⁰ cfu (unités formant des colonies) d'une culture de la souche IM2394. L'adsorption est poursuivie une nuit à 4°C, sous agitation. Les bactéries sont ensuite éliminées par centrifugation. Le surnageant est récupéré puis soumis à nouveau à 2
20 opérations d'adsorption successives comme précédemment décrit.

1C - Détermination des séquences peptidiques permettant l'identification des fragments d'ADN.

25 Des fractions aliquotes du matériel obtenu en 1A sont séchées puis resolubilisées dans le tampon de Laemmli deux fois concentré (Tris 65mM, SDS 3 %, glycérol 10 %, 2-mercaptoéthanol 5 %). On ajoute un volume d'eau équivalent.

30 Après sonication, le matériel est chauffé à 90°C pendant 2 minutes, puis soumis à une électrophorèse sur gel de polyacrylamide. Les sous-unités ainsi séparées sont transférées sur membrane PVDF (Immobilon, Millipore) pendant 16 heures à 400 mA en tampon Tris borate 50 mM, pH 8,3. Les sous-unités électrotransférées sont colorées à l'amido black et les bandes
35 correspondant à Tbp1 et Tbp2 sont récupérées et soumises au microséquençage de l'extrémité N-terminale.

Ceci est répété plusieurs fois pour établir les séquences consensus N-terminales suivantes :

5 Tbp1 IM2394 : EXVQAEQAQEKQLDTIQV
 Tbp2 IM2394 : XLXXXXSFDLDSVEXVQXXMX

(X = acide aminé non-déterminé).

10 Afin de séquencer des régions internes de Tbp2, la protéine sur membrane PVDF est soumise à digestion par la trypsine en tampon Tris 0,1 M pH 8,2. Après 4 heures de réaction à 37°C, les peptides sont extraits par de l'acide formique 70 % puis par de l'acide trifluoroacétique (TFA) 0,1 %. Ces peptides sont ensuite séparés par HPLC.

15 Pour Tbp2 IM2394, les séquences internes qui ont été établies sont les suivantes :

 S1122 : NNIVLFGPDGYLYYK
 S1125 : YTIQA
 S⁷⁷⁰ : DGENAAGPATEXVIDAYR
 20 S⁷⁶⁶ : XQIDSFQDVK
 S1126 : AAFXXXI
 S⁷⁶⁹ : XNXXXMFLQGVR
 S⁷⁷¹ : TPVSDVAAR
 S⁷⁶⁷ : XSPAFT
 25 S⁷⁶² : NAIEMGGSFXFPGNAPEG(K)
 S1128 : XQPESQQDVSENX

1D - Préparation de l'ADN génomique.

30 Le culot bactérien obtenu en 1A est resuspendu dans environ 25 ml de solution A (Tris HCl 25 mM, pH 8 contenant 50 mM de glucose et 10 mM d'EDTA) additionnée de 10 mg de protéinase K. Le mélange est laissé 10 minutes à température ambiante.

Puis on ajoute 12,5 ml de solution A contenant 10 mg de lysosyme. Une nouvelle fois, le mélange est laissé 10 minutes à température ambiante. On complète alors par 0,5 ml de sarkosyl 10 %. Le mélange est incubé 10 minutes à +4°C.

5 On ajoute ensuite 2 mg de RNase et on laisse l'incubation se poursuivre 90 minutes à 37°C. L'ADN est purifié par quatre extractions phénoliques successives. L'ADN présent dans la dernière phase aqueuse est précipité par l'éthanol. L'ADN de haut poids moléculaire est obtenu par
10 séparation sur gradient de CsCl.

1E - Clonage.

Une première banque d'ADN a été réalisée dans le vecteur lambda
15 ZAP (Figure 1), comme suit :

Une préparation d'ADN génomique a été fragmentée aux ultrasons. Les extrémités des fragments ainsi obtenus ont été rendues franches par traitement à la T₄ polymérase. Les fragments ont été méthylés. Après
20 méthylation, les fragments ont été liés à des adaptateurs *EcoRI*, traités par *EcoRI* puis insérés dans le site *EcoRI* du phage lambda ZAP II (Stratagène).

La souche *E. coli* XL1-Blue (Stratagène) a été infectée avec la banque d'ADN ainsi préparée. Les plages de lyse blanches (présence de
25 phages recombinants) ont été testées à l'aide d'un antisérum spécifique du récepteur de la transferrine de la souche IM2394 préparé tel que décrit en 1B. Ceci a permis d'identifier deux clones lambda ZAP II. Les plasmides pBluescript contenus dans ces clones ont été excisés par co-infection avec le phage-"helper" et ont été appelés pBMT1 et pBMT2.

30 Les plasmides pBMT1 et pBMT2 contiennent chacun un fragment *EcoRI* - *EcoRI* respectivement de 3,8 kb et 1,3 kb. Ils sont présentés dans la Figure 3.

35 Le séquençage de l'insert *EcoRI* - *EcoRI* de pBMT1 a été mis en oeuvre selon la méthode de shotgun (Bankier et Barrell, Biochemistry (1983))

B5 : 508), comme suit :

5 L'insert *EcoRI* - *EcoRI* de pBMT1 a été purifié puis fragmenté aux ultra-sons. Les extrémités des fragments ainsi obtenus ont été rendues franches par traitement à la T₄ polymérase. Les fragments ainsi traités ont été introduits dans un site du phage M13TG131 (décrit dans Kieny et al, Gene (1983) 26 : 91). Environ 200 clones issus de cette préparation ont été séquencés. L'analyse de ces séquences par ordinateur a permis de reconstituer la séquence complète de l'insert *EcoRI* - *EcoRI* de pBMT1.

10

La séquence codant pour l'extrémité N-terminale de Tbp1 a été localisée comme le montre la Figure 3. Compte tenu de la masse moléculaire de Tbp1, il était clair que cet insert ne comportait pas le fragment d'ADN complet codant pour Tbp1. En amont de l'extrémité 5' du gène *tbp1*, on a mis en évidence un cadre de lecture ouvert mais il n'a pas été possible d'identifier clairement une région codant pour l'extrémité N-terminale du gène *tbp2*.

15

20 Le microséquençage de régions internes de Tbp2 a donc été entrepris comme reporté précédemment en 1C. Les séquences internes qui étaient localisées vers l'extrémité C-terminale, correspondaient bien à la partie 3' du cadre de lecture ouvert en amont de *tbp1*.

25

D'autre part, l'ADN génomique de la souche IM2394, préalablement digéré par *HincII* a été analysé par Southern blot à l'aide d'une sonde d'ADN radioactive correspondant à la zone *HincII* - *HincII* de 1,5 kb de l'insert de 3,8 kb de pBMT1 ; deux bandes ont été ainsi révélées. Ceci a permis de démontrer que l'insert porté par pBMT1 résultait d'un assemblage artéfactuel de séquences issues de deux loci distincts. La séquence 5' de *tbp2* était donc absente.

30

35 La banque d'ADN génomique en lambda ZAP précédemment décrite a été criblée de nouveau ; cette fois-ci en utilisant l'insert *EcoRI* - *EcoRI* de pBMT2 comme sonde. 29 candidats ont été retenus parmi environ 200 000 plages testées. Seul le plasmide dérivé pTG2749 semblait posséder un insert nouveau par rapport à pBMT1 et pBMT2. L'insert de pTG2749 est

tel que représenté dans la Figure 3. La région de l'insert en amont du site *EcoRV* (région *EcoRV* - *EcoRI*) a été sous-clonée dans M13TG131 et séquencée par la méthode de Sanger et al, PNAS (1977) 74 : 5463 à l'aide de primers synthétiques. La séquence correspondant à l'extrémité N-terminale de Tbp2 a été ainsi retrouvée.

La séquence du fragment d'ADN codant pour Tbp2 de la souche IM2394 est présentée dans le SEQ ID NO : 1 ainsi que la séquence d'acides aminés correspondante.

Juste en amont de la séquence codant pour Tbp2 mature, l'insert de pTG2749 comporte une région génomique distincte issue d'un autre locus. La aussi, il s'agit d'une artéfact de clonage analogue à celui mis en évidence dans le cas de pBMT1.

Compte tenu des réarrangements observés et de l'absence de séquences 3' de *tbp1* et 5' de *tbp2*, la banque d'ADN génomique construite en lambda ZAP a été jugée inadaptée pour la poursuite du clonage.

Une deuxième banque d'ADN génomique a donc été construite dans un plasmide à faible nombre de copies, comme suit : une préparation d'ADN génomique a été partiellement digérée par *Sau3A*. Des fragments d'ADN d'environ 4 à 6 kb ont été purifiés après fractionnement en gradient de sucrose et insérés dans le site *BamHI* du plasmide pTG1265. Cette préparation plasmidique a servi à transformer la souche d'*E. coli* 5K. On a estimé que cette banque contenait environ 18 000 clones indépendants.

Environ 50 000 clones de la deuxième banque ont été testés à l'aide d'une sonde radioactive correspondant à l'insert *EcoRI* - *EcoRI* de pBMT2. Un seul clone a été révélé ; soit le plasmide pTG2759 qui possède un insert de 1,8 kb. La taille de cette insert a été jugée insuffisante pour contenir le gène complet codant pour Tbp1.

Une troisième banque d'ADN a été construite selon la méthode décrite au paragraphe précédent à l'exception de la souche d'*E. coli* 5K qui a été remplacée par la souche d'*E. coli* SURE (Stratagène). On a estimé que

cette banque contenait environ 60 000 clones indépendants.

5 Environ 70 000 clones de la troisième banque d'ADN ont été testés à l'aide d'une sonde radioactive correspondant au fragment *MluI* - *HincII* de 2,4 kb issu de l'insert de pTG2754 décrit dans l'Exemple 2 ci-après et représenté dans la Figure 4. Deux clones ont été révélés, soient les plasmides pTG2780 et pTG2781, représentés dans la Figure 3.

10 La séquence des inserts de pTG2780 et pTG2781 a été établie selon la méthode de Sanger. Elle est présentée dans le SEQ ID NO : 2 ainsi que la séquence d'acides aminés correspondante.

15 Une quatrième banque a été construite. Le DNA génomique a été digéré par *Sau3A* et une fraction contenant des fragments d'environ 7 kb a été purifiée sur gradient de sucrose. Cette fraction contenait un fragment correspondant au locus *tbp1,2* car elle était reconnue par une sonde d'ADN spécifique de *tbp2*. Après digestion par *EcoRV* et *XbaI* et ligation à pTG1265 digéré par *SmaI* et *XbaI*, *E. coli* 5K a été transformée. Un criblage des clones à l'aide d'une sonde spécifique de *tbp2* a été réalisé. Parmi une série de
20 clones positifs, le plasmide pTG3791 a été étudié en particulier et s'est avéré contenir des séquences 5' *tbp2* incluant la séquence codant pour le peptide signal putatif de Tbp2.

25 **EXEMPLE 2 :** Clonage des fragments d'ADN codant pour les sous-unités Tbp1 et Tbp2 du récepteur de la transferrine de la souche IM2169.

30 2A - La culture de la souche IM2169 et la purification du récepteur de la transferrine ont été effectuées dans des conditions identiques à celles décrites dans l'Exemple 1A.

2B - La préparation d'un antisérum anti-récepteur de la souche IM2169 a été réalisée selon le protocole décrit dans l'Exemple 1B.

35 2C - Les séquences peptidiques permettant l'identification des fragments d'ADN ont été déterminées selon la méthode reportée dans l'Exemple 1C.

Les microséquences qui ont été établies sont les suivantes.

Séquence consensus de l'extrémité N-terminale de Tbp1 :

ENVQAGQAQEKQLXXIQVX

5

Séquences des peptides internes de Tbp1 :

S1031 : XLS(E,W)NAGXVLXPADX

S1032 : QLDTIQVK

S1033 : TAGSSGAINEIEYENXX

10

S1034 : YVTWENVDXXXXXX

Séquence consensus de l'extrémité N-terminale de Tbp2 :

SLVXAXSFDLXSV

15

Séquences des peptides internes de Tbp2 :

S1037 : XXDNLSNAX

S1035 : XGDDGYIFYXGEKPX

S1036 : XQGXYGFAMX

S1040 : XQATGHENFQYVYSGXFYK

20

2D - La préparation de l'ADN génomique de la souche IM2169 a été réalisée selon le protocole décrit dans l'Exemple 1D.

2E - Clonage

25

Une première banque d'ADN génomique (fragments d'ADN *Sau3A* partiel ; pTG1265 ; *E. coli* 5K) a été construite comme précédemment décrit dans l'Exemple 1. On a estimé que cette banque contenait environ 40 000 clones indépendants, dont environ 70 % possédaient un insert de 4-6 kb.

30

130 000 clones de cette banque ont été testés à l'aide d'une sonde radioactive correspondant à l'insert *EcoRI* - *EcoRI* de pBMT2. 42 clones ont été analysés, parmi lesquels 2 ont été retenus : les plasmides pTG2753 et pTG2754 qui sont tels que montrés dans la Figure 4. Les analyses en Southern blot ont montré que les cartes de restriction des inserts de pTG2753 et pTG2754 correspondaient à la carte de restriction de l'ADN

35

génomique.

La détermination des séquences nucléotidiques et la recherche des régions codant pour les extrémités N-terminales et les régions internes ont
5 démontré que :

- l'insert de 1,9 kb de pTG2753 contient la partie 3' du gène *tbp2* et la partie 5' du gène *tbp1* ; et
- 10 - l'insert de pTG2754 contient la partie 3' du gène *tbp2* et les parties 5' et 3' du gène *tbp1*, en rupture de phase.

Cette première banque n'a donc pas permis de cloner des fragments d'ADN complets codant pour Tbp1 ou Tbp2.
15

Une deuxième banque génomique a été construite comme précédemment mais à partir d'ADN génomique digéré par *XbaI*. Les fragments d'ADN ont été purifiés après fractionnement en gradient de sucrose. Chaque fraction (d'environ 500 µl) a été testée par Southern blot
20 avec une sonde radioactive correspondant à l'extrémité 3' de *tbp1* (fragment de l'insert de pTG2754). La fraction présentant une réaction d'hybridation et contenant des fragments d'environ 6 kb a été clonée dans pTG1265. La souche *E. coli* 5K a été transformée.

25 Environ 2 400 clones de cette banque ont été testés à l'aide d'une sonde radioactive correspondant au fragment *HincII* - *MluI* de 0,6 kb issu de pTG2754. Cinq clones ont été caractérisés, parmi lesquels 2 ont été retenus : soient pTG3720 et pTG3721, tels que montrés dans la Figure 4, qui contiennent tous deux les gènes *tbp1* et *tbp2*.

30 Afin de compléter la séquence nucléotidique codant pour Tbp1, l'insert de pTG3720 a été séquencé dans la région où se situait la rupture de phase découverte dans l'insert de pTG2754. Ce séquençage a permis de mettre en évidence que la rupture de phase de l'insert de pTG2754 était due à une
35 délétion de 22 bp. La séquence complète du fragment d'ADN est telle que montrée dans le SEQ ID NO : 3.

Le séquençage de l'insert de pTG3720 a été poursuivi pour établir la séquence de *thp2*. Celle-ci a bien été identifiée ; mais là aussi une rupture de phase a été constatée.

5 Finalement la séquence de *thp2* a été déterminée à partir du plasmide pTG3721. Elle est telle que montrée dans le SEQ ID NO : 4.

EXEMPLE 3 : Expression du fragment d'ADN codant pour la sous-unité Tbp2 de la souche IM2394.

10

3A. Construction du vecteur d'expression pTG3786.

Le site *SphI* du plasmide para13 (Figure 5 ; Cagnon et al, Prot. Eng. (1991) 4 : 843) a été détruit par traitement à la polymérase Klenow, pour
15 donner le plasmide pTG3704. pTG3704 a été linéarisé par coupure *NcoI*, traité à la polymérase Klenow pour rendre les extrémités franches, puis digéré par *HindIII*.

20 D'autre part, on a synthétisé les oligonucléotides OTG4015 et OTG4016 que l'on a appariés.

OTG4015 : 5' AAATACCTATTGCCTACGGCAGCCGCTGGACTGTTATTACT
CGCTGCCCAACCAGCGATGGCATGCTTTCCCACGCGTTTCCCA 3'

25 OTG4016 : 5' AGCTTGGGAAAACGCGTGGGAAAGCATGCCATCGCTGGTTGGGCA
GCGAGTAATAACAGTCCAGCGGCTGCCGTAGGCAATAGGTATTT 3'

30 Le fragment d'ADN double brin OTG4015/OTG4016 a été inséré dans para13 traité comme précédemment décrit, pour donner le plasmide pTG3717 dans lequel on avait reconstitué la séquence codant pour la partie N-terminale du précurseur de la protéine PelB d'*Erwinia carotovora* (Lei et al, J. Bact. (1987) 169 : 4379) ; Soit :

35 ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA CTG
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu

SphI

TTA TTA CTC GCT GCC CAA CCA GCG ATG GCA TGCTTT
Leu Leu Leu Ala Ala Gln Pro Ala Met Ala

5 *MluI* *HindIII*
CCCACGCGTTTTCCCA AGCTT.....

(en souligné, apparaissent les extrémités de pTG3704)

10 A partir du plasmide pTG2749, on a généré par PCR à l'aide des amorces OTG4011 et OTG 4012, un fragment incluant la région codant pour la partie N-terminale de Tbp2, jusqu'au site *MluI* interne, tel que montré dans la Figure 6.

15 OTG4011 :
 BamHI *SphI*
5' AAAAAGGATCC/GCA TGC CTG GGT GGC GGC GGC AGT TTC 3'
 Cys Leu Gly

20 OTG4012 :
 BamHI *MluI*
5' AAAAGGATCCG AAT GGT GTA ACG CGT AGT TTT TAT 3'

25 Le fragment généré par PCR a été digéré par *BamHI*, puis inséré dans le site *BamHI* du phage M13TG131, pour donner M13TG3724. La séquence de ce fragment a été vérifiée par séquençage.

30 A partir de M13TG3724, on a récupéré la région codant pour la partie N-terminale de Tbp2 sous forme d'un fragment *SphI* - *MluI* que l'on insert dans pTG3717 préalablement digéré par *SphI* et *MluI*, pour donner le plasmide pTG3743.

35 A partir du plasmide pBMT1, on a récupéré la région codant pour la partie C-terminale de Tbp2 sous forme d'un fragment *MluI* - *BanI* dont l'extrémité cohésive *BanI* avait été rendue franche par traitement à la polymérase Klenow. On a inséré ce fragment dans pTG3743 préalablement

- 26 -

digéré par *HindIII*, traité à la polymérase Klenow et finalement digéré par *MluI*. On obtient ainsi le plasmide pTG3786.

3B. Production de la sous-unité Tbp2.

5

E. coli MC1061 (Casadaban & Cohen, J. Mol. Biol. (1980) 138 : 179) est transformée par pTG3786 puis mise en culture à 37°C, en milieu LB supplémenté avec 2 g/l de glycérol. A la culture est en phase exponentielle, on ajoute 0,2 g/l d'arabinose. L'incubation a été poursuivie durant 6 hr supplémentaires. L'expression a été observée moins d'une heure après l'addition d'arabinose.

10

L'électrophorèse sur gel d'acrylamide d'un échantillon du lysat cellulaire total a mis en évidence la présence d'une protéine d'environ 70 kD qui est capable de fixer la transferrine humaine marquée à la peroxydase.

15

EXEMPLE 4 : Expression du fragment d'ADN codant pour la sous-unité *tbpI* de la souche IM2169.

20

4A. Construction du vecteur d'expression pTG37796.

Un fragment synthétique constitué des oligonucléotides OTG4038 et OTG4039 préalablement appariés, a été inséré dans le plasmide pTG3704 digéré par *NcoI* et *HindIII*, générant ainsi le plasmide pTG3756.

25

OTG4038 :

5' CATGGCTGCAGGRACCACGCGTGAATTCCCCGGGTCTAGA 3'

OTG4039 :

30

5' AGCTTCTAGACCCGGGGAATTCACGCGTGGTACCTGCAGC 3'

A partir du plasmide pTG2754, on a généré par PCR à l'aide des amorces OTG4037 et OTG4014 un fragment incluant la région codant pour l'extrémité N-terminale du précurseur de Tbp1 jusqu'au site *MluI*.

OTG4037 :

5' TTTCCCGGATCCGC ATG CAA CAG CAA CAT TTG TTC CGA TTA 3

*Bam*HI *Sph*I

Met Gln Gln Gln...

5

OTG4014 :

5' AAAAGGATCCGGGGTCGTAACGCGTCAGGTCGCGG 3'

*Bam*HI

*Mlu*I

10

Ce fragment PCR a été digéré par *Bam*HI et cloné dans le site *Bam*HI de M13TG131 pour générer M13TG3738. La séquence de ce fragment a été vérifiée.

15

M13TG3738 a ensuite été linéarisé par *Sph*I, traité à la T4 DNA polymérase pour rendre les extrémités franches, puis digéré par *Mlu*I afin d'isoler le fragment porteur de la région codant pour l'extrémité N-terminale du précurseur de Tbp1.

20

Ce fragment a été inséré dans pTG3756 digéré par *Nco*I, traité à la T4 DNA polymérase puis digéré par *Mlu*I, pour générer le plasmide pTG3778. La séquence de la jonction *Nco*I^r/*Sph*I^r a été vérifiée.

25

Le fragment *Mlu*I - *Xba*I de pTG3720 codant pour la majeure partie de Tbp1 (3'*tbp1*) a été inséré dans le plasmide pTG3778. Le plasmide final ainsi obtenu est le plasmide pTG3779.

4B. Production de la sous-unité Tbp1.

30

E. coli MC1061 a été transformé par pTG3779 puis mise en culture à 37°C en milieu LB. A la culture en phase exponentielle, on a ajouté 0,2 g/l d'arabinose. L'incubation a été poursuivie durant 4 heures.

35

L'électrophorèse sur gel d'acrylamide d'un échantillon du lysat cellulaire total a mis en évidence la présence d'une protéine d'environ 100 Kd qui est capable de fixer la transferrine humaine marquée à la peroxydase.

SEQ ID NO : 1

Objet : Séquence de l'ADN génomique de la souche de *N. meningitidis* IM2394 codant pour la sous unité Tbp2 et séquence protéique déduite. En gras sont indiqués le peptide signal et le site *MluI*

										<u>ATG</u>	AAC	AAT	CCA	TTG	15
										Met	Asn	Asn	Pro	Leu	
										-20					
GTA	AAT	CAG	GCT	GCT	ATG	GTG	CTG	CCT	GTG	TTT	TTG	TTG	AGT	GCT	60
Val	Asn	Gln	Ala	Ala	Met	Val	Leu	Pro	Val	Phe	Leu	Leu	Ser	Ala	
-15					-10					-5				-1	
TGT	CTG	GGT	GGC	GGC	GGC	AGT	TTC	GAT	TTG	GAC	AGC	GTG	GAA	ACC	105
Cys	Leu	Gly	Gly	Gly	Gly	Ser	Phe	Asp	Leu	Asp	Ser	Val	Glu	Thr	
1				5					10					15	
GTG	CAA	GAT	ATG	CAC	TCC	AAA	CCT	AAG	TAT	GAG	GAT	GAA	AAA	AGC	130
Val	Gln	Asp	Met	His	Ser	Lys	Pro	Lys	Tyr	Glu	Asp	Glu	Lys	Ser	
				20					25					30	
CAG	CCT	GAA	AGC	CAA	CAG	GAT	GTA	TCG	GAA	AAC	AGC	GGC	GCG	GCT	195
Gln	Pro	Glu	Ser	Gln	Gln	Asp	Val	Ser	Glu	Asn	Ser	Gly	Ala	Ala	
				35					40					45	
TAT	GGC	TTT	GCA	GTA	AAA	CTA	CCT	CGC	CGG	AAT	GCA	CAT	TTT	AAT	240
Tyr	Gly	Phe	Ala	Val	Lys	Leu	Pro	Arg	Arg	Asn	Ala	His	Phe	Asn	
				50					55					60	
CCT	AAA	TAT	AAG	GAA	AAG	CAC	AAA	CCA	TTG	GGT	TCA	ATG	GAT	TGG	285
Pro	Lys	Tyr	Lys	Glu	Lys	His	Lys	Pro	Leu	Gly	Ser	Met	Asp	Trp	
				65					70					75	
AAA	AAA	CTG	CAA	AGA	GGA	GAA	CCA	AAT	AGT	TTT	AGT	GAG	AGG	GAT	330
Lys	Lys	Leu	Gln	Arg	Gly	Glu	Pro	Asn	Ser	Phe	Ser	Glu	Arg	Asp	
				80					85					90	
GAA	TTG	GAA	AAA	AAA	CGG	GGT	AGT	TCT	GAA	CTT	ATT	GAA	TCA	AAA	375
Glu	Leu	Glu	Lys	Lys	Arg	Gly	Ser	Ser	Glu	Leu	Ile	Glu	Ser	Lys	
				95					100					105	
TGG	GAA	GAT	GGG	CAA	AGT	CGT	GTA	GTT	GGT	TAT	ACA	AAT	TTC	ACT	420
Trp	Glu	Asp	Gly	Gln	Ser	Arg	Val	Val	Gly	Tyr	Thr	Asn	Phe	Thr	
				110					115					120	
TAT	GTC	CGT	TCG	GGA	TAT	GTT	TAC	CTT	AAT	AAA	AAT	AAT	ATT	GAT	465
Tyr	Val	Arg	Ser	Gly	Tyr	Val	Tyr	Leu	Asn	Lys	Asn	Asn	Ile	Asp	
				125					130					135	
ATT	AAG	AAT	AAT	ATA	GTT	CTT	TTT	GGA	CCT	GAC	GGA	TAT	CTT	TAC	510
Ile	Lys	Asn	Asn	Ile	Val	Leu	Phe	Gly	Pro	Asp	Gly	Tyr	Leu	Tyr	
				140					145					150	

TAT AAA GGG AAA GAA CCT TCC AAG GAG CTG CCA TCG GAA AAG ATA	555
Tyr Lys Gly Lys Glu Pro Ser Lys Glu Leu Pro Ser Glu Lys Ile	155 160 165
ACT TAT AAA GGT ACT TGG GAT TAT GTT ACT GAT GCT ATG GAA AAA	600
Thr Tyr Lys Gly Thr Trp Asp Tyr Val Thr Asp Ala Met Glu Lys	170 175 180
CAA AGG TTT GAA GGA TTG GGT AGT GCA GCA GGA GGA GAT AAA TCG	645
Gln Arg Phe Glu Gly Leu Gly Ser Ala Ala Gly Gly Asp Lys Ser	185 190 195
GGG CCG TTG TCT GCA TTA GAA GAA GGG GTA TTG CGT AAT CAG GCA	690
Gly Ala Leu Ser Ala Leu Glu Glu Gly Val Leu Arg Asn Gln Ala	200 205 210
GAG GCA TCA TCC GGT CAT ACC GAT TTT GGT ATG ACT AGT GAG TTT	735
Glu Ala Ser Ser Gly His Thr Asp Phe Gly Met Thr Ser Glu Phe	215 220 225
GAG GTT GAT TTT TCT GAT AAA ACA ATA AAG GGC ACA CTT TAT CGT	780
Glu Val Asp Phe Ser Asp Lys Thr Ile Lys Gly Thr Leu Tyr Arg	230 235 240
AAC AAC CGT ATT ACT CAA AAT AAT AGT GAA AAC AAA CAA ATA AAA	825
Asn Asn Arg Ile Thr Gln Asn Asn Ser Glu Asn Lys Gln Ile Lys	245 250 255
ACT ACG CGT TAC ACC ATT CAA GCA ACT CTT CAC GGC AAC CGT TTC	870
Thr Thr Arg Tyr Thr Ile Gln Ala Thr Leu His Gly Asn Arg Phe	260 265 270
AAA GGT AAG GCG TTG GCG GCA GAT AAA GGT GCA ACA AAT GGA AGT	915
Lys Gly Lys Ala Leu Ala Ala Asp Lys Gly Ala Thr Asn Gly Ser	275 280 285
CAT CCC TTT ATT TCC GAC TCC GAC AGT TTG GAA GGC GGA TTT TAC	960
His Pro Phe Ile Ser Asp Ser Asp Ser Leu Glu Gly Gly Phe Tyr	290 295 300
GGG CCG AAA GGC GAG GAA CTT GCC GGT AAA TTC TTG AGC AAC GAC	1005
Gly Pro Lys Gly Glu Glu Leu Ala Gly Lys Phe Leu Ser Asn Asp	305 310 315
AAC AAA GTT GCA GCG GTG TTT GGT GCG AAG CAG AAA GAT AAG AAG	1050
Asn Lys Val Ala Ala Val Phe Gly Ala Lys Gln Lys Asp Lys Lys	320 325 330
GAT GGG GAA AAC GCG GCA GGG CCT GCA ACG GAA ACC GTG ATA GAT	1095
Asp Gly Glu Asn Ala Ala Gly Pro Ala Thr Glu Thr Val Ile Asp	335 340 345
GCA TAC CGT ATT ACC GGC GAG GAG TTT AAG AAA GAG CAA ATA GAC	1140
Ala Tyr Arg Ile Thr Gly Glu Glu Phe Lys Lys Glu Gln Ile Asp	350 355 360
AGT TTT GGA GAT GTG AAA AAG CTG CTG GTT GAC GGA GTG GAG CTT	1185
Ser Phe Gly Asp Val Lys Lys Leu Leu Val Asp Gly Val Glu Leu	365 370 375
TCA CTG CTG CCG TCT GAG GGC AAT AAG GCG GCA TTT CAG CAC GAG	1230
Ser Leu Leu Pro Ser Glu Gly Asn Lys Ala Ala Phe Gln His Glu	380 385 390

- 30 -

ATT GAG CAA AAC GGC GTG AAG GCA ACG GTG TGT TGT TCC AAC TTG	1275
Ile Glu Gln Asn Gly Val Lys Ala Thr Val Cys Cys Ser Asn Leu	
395 400 405	
GAT TAC ATG AGT TTT GGG AAG CTG TCA AAA GAA AAT AAA GAC GAT	1320
Asp Tyr Met Ser Phe Gly Lys Leu Ser Lys Gku Asn Lys Asp Asp	
410 415 420	
ATG TTC CTG CAA GGT GTC CGC ACT CCA GTA TCC GAT GTG GCG GCA	1365
Met Phe Leu Gln Gly Val Arg Thr Pro Val Ser Asp Val Ala Ala	
425 430 435	
AGG ACG GAG GCA AAC GCC AAA TAT CGC GGT ACT TGG TAC GGA TAT	1410
Arg Thr Glu Ala Lys Tyr Arg Gly Thr Gly Thr Trp Tyr Gly Tyr	
440 445 450	
ATT GCC AAC GGC ACA AGC TGG AGC GGC GAA GCC TCC AAT CAG GAA	1455
Ile Ala Asn Gly Thr Ser Trp Ser Gly Glu Ala Ser Asn Gln Glu	
455 460 465	
GGT GGT AAT AGG GCA GAG TTT GAC GTG GAT TTT TCC ACT AAA AAA	1500
Gly Gly Asn Arg Ala Glu Phe Asp Val Asp Phe Ser Thr Lys Lys	
470 475 480	
ATC AGT GGC ACA CTG ACG GCA AAA GAC CGT ACG TCT CCT GCG TTT	1545
Ile Ser Gly Thr Leu Thr Ala Lys Asp Arg Thr Ser Pro Ala Phe	
485 490 495	
ACT ATT ACT GCC ATG ATT AAG GAC AAC GGT TTT TCA GGT GTG GCG	1590
Thr Ile Thr Ala Met Ile Lys Asp Asn Gly Phe Ser Gly Val Ala	
500 505 510	
AAA ACC GGT GAA AAC GGC TTT GCG CTG GAT CCG CAA AAT ACC GGA	1635
Lys Thr Gly Glu Asn Gly Phe Ala Leu Asp Pro Gln Asn Thr Gly	
515 520 525	
AAT TCC CAC TAT ACG CAT ATT GAA GCC ACT GTA TCC GGC GGT TTC	1680
Asn Ser His Tyr Thr His Ile Glu Ala Thr Val Ser Gly Gly Phe	
530 535 540	
TAC GGC AAA AAC GCC ATC GAG ATG GGC GGA TCG TTC TCA TTT CCG	1725
Tyr Gly Lys Asn Ala Ile Glu Met Gly Gly Ser Phe Ser Phe Pro	
545 550 555	
GGA AAT GCA CCA GAG GGA AAA CAA GAA AAA GCA TCG GTG GTA TTC	1770
Gly Asn Ala Pro Glu Gly Lys Gln Glu Lys Ala Ser Val Val Phe	
560 565 570	
GGT GCG AAA CGC CAA CAG CTT GTG CAA TAA GCACGGCT	1808
Gly Ala Lys Arg Gln Gln Leu Val Gln	
575	

SEO ID NO : 2

Objet : Séquence de l'ADN génomique de la souche de *N. meningitidis* IM2394 codant pour le précurseur de la sous-unité Tbp1 du récepteur de la transferrine et séquence protéique déduite. Le peptide signal est indiqué en caractères gras.

CTTCCGATG CCGTCTGAAA GCGAAGATTA GGGAAACACT															40
ATG CAA CAG CAA CAT TTG TTC CGA TTA AAT ATT TTA TGC CTG TCT	85														
Met Gln Gln Gln His Leu Phe Arg Leu Asn Ile Leu Cys Leu Ser															
-24 -20 -15 -10															
TTA ATG ACC GCG CTG CCC GTT TAT GCA GAA AAT GTG CAA GCC GAA	130														
Leu Met Thr Ala Leu Pro Val Tyr Ala Glu Asn Val Gln Ala Glu															
-5 -1 1 5															
CAA GCA CAG GAA AAA CAG TTG GAT ACC ATA CAG GTA AAA GCC AAA	175														
Gln Ala Gln Glu Lys Gln Leu Asp Thr Ile Gln Val Lys Ala Lys															
10 15 20															
AAA CAG AAA ACC CGC CGC GAT AAC GAA GTA ACC GGG CTG GGC AAG	220														
Lys Gln Lys Thr Arg Arg Asp Asn Glu Val Thr Gly Leu Gly Lys															
25 30 35															
TTG GTC AAG TCT TCC GAT ACG CTA AGT AAA GAA CAG GTT TTG AAT	265														
Leu Val Lys Ser Ser Asp Thr Leu Ser Lys Glu Gln Val Leu Asn															
40 45 50															
ATC CGA GAC CTG ACC CGT TAT GAT CCG GGT ATT GCC GTG GTC GAA	310														
Ile Arg Asp Leu Thr Arg Tyr Asp Pro Gly Ile Ala Val Val Glu															
55 60 65															
CAG GGT CGG GGC GCA AGT TCC GGC TAT TCA ATA CGC GGC ATG GAT	355														
Gln Gly Arg Gly Ala Ser Ser Gly Tyr Ser Ile Arg Gly Met Asp															
70 75 80															
AAA AAC CGC GTT TCC TTA ACG GTA GAC GGC GTT TCG CAA ATA CAG	400														
Lys Asn Arg Val Ser Leu Thr Val Asp Gly Val Ser Gln Ile Gln															
85 90 95															
TCC TAC ACC GCG CAG GCG GCA TTG GGT GGG ACG AGG ACG GCG GGT	445														
Ser Tyr Thr Ala Gln Ala Ala Leu Gly Gly Thr Arg Thr Ala Gly															
100 105 110															
AGC AGC GGC GCA ATC AAT GAA ATC GAG TAT GAA AAC GTC AAG GCC	490														
Ser Ser Gly Ala Ile Asn Glu Ile Glu Tyr Glu Asn Val Lys Ala															
115 120 125															
GTT GAA ATC AGC AAG GGT TCG AAT TCA TCA GAA TAC GGA AAC GGC	535														
Val Glu Ile Ser Lys Gly Ser Asn Ser Ser Glu Tyr Gly Asn Gly															
130 135 140															
GCA TTG GCA GGT TCG GTC GCA TTT CAA ACC AAA ACC GCA GCC GAC	580														
Ala Leu Ala Gly Ser Val Ala Phe Gln Thr Lys Thr Ala Ala Asp															
145 150 155															

ATT ATC GGA GAG GGA AAA CAG TGG GGC ATT CAG AGT AAA ACT GCC Ile Ile Gly Glu Gly Lys Gln Trp Gly Ile Gln Ser Lys Thr Ala 160 165 170	625
TAT TCG GGA AAA GAC CAT GCC CTG ACG CAA TCC CTT GCG CTT GCC Tyr Ser Gly Lys Asp His Ala Leu Thr Gln Ser Leu Ala Leu Ala 175 180 185	670
GGA CGC AGC GGC GGC GCG GAA GCC CTC CTT ATT TAT ACT AAA CGG Gly Arg Ser Gly Gly Ala Glu Ala Leu Leu Ile Tyr Thr Lys Arg 190 195 200	715
CGG GGT CGG GAA ATC CAT GCG CAT AAA GAT GCC GGC AAG GGT GTG Arg Gly Arg Glu Ile His Ala His Lys Asp Ala Gly Lys Gly Val 205 210 215	760
CAG AGC TTC AAC CGG CTG GTG TTG GAC GAG GAC AAG AAG GAG GGT Gln Ser Phe Asn Arg Leu Val Leu Asp Glu Asp Lys Lys Glu Gly 220 225 230	805
GGC AGT CAG TAC AGA TAT TTC ATT GTC GAA GAA GAA TGC CAC AAT Gly Ser Gln Tyr Arg Tyr Phe Ile Val Glu Glu Glu Cys His Asn 235 240 245	850
GGA TAT GCG GCC TGT AAA AAC AAG CTG AAA GAA GAT GCC TCG GTC Gly Tyr Ala Ala Cys Lys Asn Lys Leu Lys Glu Asp Ala Ser Val 250 255 260	895
AAA GAT GAG CGC AAA ACC GTC AGC ACG CAG GAT TAT ACC GGC TCC Lys Asp Glu Arg Lys Thr Val Ser Thr Gln Asp Tyr Thr Gly Ser 265 270 275	940
AAC CGC TTA CTT GCG AAC CCG CTT GAG TAT GGC AGC CAA TCA TGG Asn Arg Leu Leu Ala Asn Pro Leu Glu Tyr Gly Ser Gln Ser Trp 280 285 290	985
CTG TTC CGA CCG GGT TGG CAT TTG GAC AAC CGC CAT TAT GTC GGA Leu Phe Arg Pro Gly Trp His Leu Asp Asn Arg His Tyr Val Gly 295 300 305	1030
GCC GTT CTC GAA CGT ACG CAG CAG ACC TTT GAT ACA CGG GAT ATG Ala Val Leu Glu Arg Thr Gln Gln Thr Phe Asp Thr Arg Asp Met 310 315 320	1075
ACT GTT CCT GCC TAT TTT ACC AGT GAA GAT TAT GTA CCC GGT TCG Thr Val Pro Ala Tyr Phe Thr Ser Glu Asp Tyr Val Pro Gly Ser 325 330 335	1120
CTG AAA GGT CTT GGC AAA TAT TCG GGC GAT AAT AAG GCA GAA AGG Leu Lys Gly Leu Gly Lys Tyr Ser Gly Asp Asn Lys Ala Glu Arg 340 345 350	1165
CTG TTT GTT CAG GGA GAG GGC AGT ACA TTG CAG GGT ATC GGT TAC Leu Phe Val Gln Gly Glu Gly Ser Thr Leu Gln Gly Ile Gly Tyr 355 360 365	1210
GGT ACC GGC GTG TTT TAT GAT GAA CGC CAT ACT AAA AAC CGC TAC Gly Thr Gly Val Phe Tyr Asp Glu Arg His Thr Lys Asn Arg Tyr 370 375 380	1255
GGG GTC GAA TAT GTT TAC CAT AAT GCT GAT AAG GAT ACC TGG GCC Gly Val Glu Tyr Val Tyr His Asn Ala Asp Lys Asp Thr Trp Ala 385 390 395	1300

GAT TAC GCC CGA CTT TCT TAT GAC CGG CAA GGT ATA GAT TTG GAC Asp Tyr Ala Arg Leu Ser Tyr Asp Arg Gln Gly Ile Asp Leu Asp 400 405 410	1345
AAC CGT TTG CAG CAG ACG CAT TGC TCT CAC GAC GGT TCG GAT AAA Asn Arg Leu Gln Gln Thr His Cys Ser His Asp Gly Ser Asp Lys 415 420 425	1390
AAT TGC CGT CCC GAC GGC AAT AAA CCG TAT TCT TTC TAT AAA TCC Asn Cys Arg Pro Asp Gly Asn Lys Pro Tyr Ser Phe Tyr Lys Ser 430 435 440	1435
GAC CGG ATG ATT TAT GAA GAA AGC CGA AAC CTG TTC CAA GCA GTA Asp Arg Met Ile Tyr Glu Glu Ser Arg Asn Leu Phe Gln Ala Val 445 450 455	1480
TTT AAA AAG GCA TTT GAT ACG GCC AAA ATC CGT CAC AAT TTG AGT Phe Lys Lys Ala Phe Asp Thr Ala Lys Ile Arg His Asn Leu Ser 460 465 470	1525
ATC AAT CTA GGG TAC GAC CGC TTT AAG TCG CAA TTG TCC CAC AGC Ile Asn Leu Gly Tyr Asp Arg Phe Lys Ser Gln Leu Ser His Ser 475 4800 485	1570
GAT TAT TAT CTT CAA AAC GCA GTT CAG GCA TAT GAT TTG ATA ACC Asp Tyr Tyr Leu Gln Asn Ala Val Gln Ala Tyr Asp Leu Ile Thr 490 495 500	1615
CCG AAA AAG CCT CCG TTT CCC AAC GGA AGC AAA GAC AAC CCG TAT Pro Lys Lys Pro Pro Phe Pro Asn Gly Ser Lys Asp Asn Pro Tyr 505 510 515	1660
AGG GTG TCT ATC GGC AAG ACC ACG GTC AAT ACA TCG CCG ATA TGC Arg Val Ser Ile Gly Lys Thr Thr Val Asn Thr Ser Pro Ile Cys 520 525 530	1705
CGT TTC GGC AAT AAC ACC TAT ACA GAC TGC ACA CCG AGG AAT ATC Arg Phe Gly Asn Asn Thr Tyr Thr Asp Cys Thr Pro Arg Asn Ile 535 540 545	1750
GGC GGC AAC GGT TAT TAT GCA GCC GTT CAA GAC AAT GTC CGT TTG Gly Gly Asn Gly Tyr Tyr Ala Ala Val Gln Asp Asn Val Arg Leu 550 555 560	1795
GGC AGG TGG GCG GAT GTC GGA GCA GGC ATA CGT TAC GAT TAC CGC Gly Arg Trp Ala Asp Val Gly Ala Gly Ile Arg Tyr Asp Tyr Arg 565 570 575	1840
AGC ACG CAT TCG GAA GAT AAG AGT GTC TCT ACC GGC ACT CAC CGC Ser Thr His Ser Glu Asp Lys Ser Val Ser Thr Gly Thr His Arg 580 585 590	1885
AAC CTT TCT TGG AAC GCG GGC GTA GTC CTC AAA CCT TTC ACC TGG Asn Leu Ser Trp Asn Ala Gly Val Val Leu Lys Pro Phe Thr Trp 595 600 605	1930
ATG GAT TTG ACT TAT CGC GCT TCT ACG GGC TTC CGT CTG CCG TCG Met Asp Leu Thr Tyr Arg Ala Ser Thr Gly Phe Arg Leu Pro Ser 610 615 620	1975
TTT GCC GAA ATG TAT GGC TGG AGA GCC GGG GAG TCT TTG AAA ACG Phe Ala Glu Met Tyr Gly Trp Arg Ala Gly Glu Ser Leu Lys Thr 625 630 635	2020

GGC GGC AAC GGT TAT TAT GCA GCC GTT CAA GAC AAT GTC CGT TTG Gly Gly Asn Gly Tyr Tyr Ala Ala Val Gln Asp Asn Val Arg Leu 550 555 560	1795
GGC AGG TGG GCG GAT GTC GGA GCA GGC ATA CGT TAC GAT TAC CGC Gly Arg Trp Ala Asp Val Gly Ala Gly Ile Arg Tyr Asp Tyr Arg 565 570 575	1840
AGC ACG CAT TCG GAA GAT AAG AGT GTC TCT ACC GGC ACT CAC CGC Ser Thr His Ser Glu Asp Lys Ser Val Ser Thr Gly Thr His Arg 580 585 590	1885
AAC CTT TCT TGG AAC GCG GGC GTA GTC CTC AAA CCT TTC ACC TGG Asn Leu Ser Trp Asn Ala Gly Val Val Leu Lys Pro Phe Thr Trp 595 600 605	1930
ATG GAT TTG ACT TAT CGC GCT TCT ACG GGC TTC CGT CTG CCG TCG Met Asp Leu Thr Tyr Arg Ala Ser Thr Gly Phe Arg Leu Pro Ser 610 615 620	1975
TTT GCC GAA ATG TAT GGC TGG AGA GCC GGG GAG TCT TTG AAA ACG Phe Ala Glu Met Tyr Gly Trp Arg Ala Gly Glu Ser Leu Lys Thr 625 630 635	2020
TTG GAT CTG AAA CCG GAA AAA TCC TTT AAT AGA GAG GCA GGT ATT Leu Asp Leu Lys Pro Glu Lys Ser Phe Asn Arg Glu Ala Gly Ile 640 645 650	2065
GTA TTT AAA GGG GAC TTC GGC AAT TTG GAA GCC AGC TAT TTC AAC Val Phe Lys Gly Asp Phe Gly Asn Leu Glu Ala Ser Tyr Phe Asn 655 660 665	2110
AAT GCC TAT CGC GAC CTG ATT GCA TTC GGT TAT GAA ACC CGA ACT Asn Ala Tyr Arg Asp Leu Ile Ala Phe Gly Tyr Glu Thr Arg Thr 670 675 680	2155
CAA AAC GGG CAA ACT TCG GCT TCT GGC GAC CCC GGA TAC CGA AAT Gln Asn Gly Gln Thr Ser Ala Ser Gly Asp Pro Gly Tyr Arg Asn 685 690 695	2200
GCC CAA AAT GCA CGG ATA GCC GGT ATC AAT ATT TTG GGT AAA ATC Ala Gln Asn Ala Arg Ile Ala Gly Ile Asn Ile Leu Gly Lys Ile 700 705 710	2245
GAT TGG CAC GGC GTA TGG GGC GGG TTG CCG GAC GGG TTG TAT TCC Asp Trp His Gly Val Trp Gly Gly Leu Pro Asp Gly Leu Tyr Ser 715 720 725	2290
ACG CTT GCC TAT AAC CGT ATC AAG GTC AAA GAT GCC GAT ATA CGC Thr Leu Ala Tyr Asn Arg Ile Lys Val Lys Asp Ala Asp Ile Arg 730 735 740	2335
GCC GAC AGG ACG TTT GTA ACT TCA TAT CTC TTT GAT GCC GTC CAA Ala Asp Arg Thr Phe Val Thr Ser Tyr Leu Phe Asp Ala Val Gln 745 750 755	2380
CCT TCA CGA TAT GTA TTG GGT TTG GGT TAC GAC CAT CCT GAC GGA Pro Ser Arg Tyr Val Leu Gly Leu Gly Tyr Asp His Pro Asp Gly 760 765 770	2425
ATA TGG GGC ATC AAT ACG ATG TTT ACT TAT TCC AAG GCA AAA TCT Ile Trp Gly Ile Asn Thr Met Phe Thr Tyr Ser Lys Ala Lys Ser 775 780 785	2470

- 35 -

ACA TTT AGC TTG GAA ATG AAG TTT TAA ACGTCCAAAC GCCGCAAATG
Thr Phe Ser Leu Glu Met Lys Phe
880

2787

CCGTCTGAAA GGCT

2801

SEQ ID NO : 3

Objet : Séquence de l'ADN génomique de la souche de *N. meningitidis* IM2169
codant pour le précurseur de la sous-unité Tbp1 et séquence protéique déduite.
Le peptide signal est indiqué en caractères gras.

	ATCAAGAATA AGGCTTCAGA	20
	CGGCATCGCT CCTTCCGATA CCGTCTGAAA GCGAAGATTA GGGAAACATT	70
ATG	CAA CAG CAA CAT TTG TTC CGA TTA AAT ATT TTA TGC CTG TCG	115
Met	Gln Gln Gln His Leu Phe Arg Leu Asn Ile Leu Cys Leu Ser	
-24	-20 -15 -10	
CTG	ATG ACT GCG CTG CCT GCT TAT GCA GAA AAT GTG CAA GCC GGA	160
Leu	Met Thr Ala Leu Pro Ala Tyr Ala Glu Asn Val Gln Ala Gly	
	-5 -1 1 5	
CAA	GCA CAG GAA AAA CAG TTG GAT ACC ATA CAG GTA AAA GCC AAA	205
Gln	Ala Gln Glu Lys Gln Leu Asp Thr Ile Gln Val Lys Ala Lys	
	10 15 20	
AAA	CAG AAA ACC CGC CGC GAT AAC GAA GTA ACC GGT CTG GGC AAA	250
Lys	Gln Lys Thr Arg Arg Asp Asn Glu Val Thr Gly Leu Gly Lys	
	25 30 35	
TTG	GTC AAA ACC GCC GAC ACC CTC AGC AAG GAA CAG GTA CTC GAT	295
Leu	Val Lys Thr Ala Asp Thr Leu Ser Lys Glu Gln Val Leu Asp	
	40 45 50	
ATC	CGC GAC CTG ACG CGT TAC GAC CCC GGC ATC GCC GTG GTC GAA	340
Ile	Arg Asp Leu Thr Arg Tyr Asp Pro Gly Ile Ala Val Val Glu	
	55 60 65	
CAG	GGG CGC GGC GCA AGT TCG GGC TAC TCG ATA CGC GGT ATG GAC	385
Gln	Gly Arg Gly Ala Ser Ser Gly Tyr Ser Ile Arg Gly Met Asp	
	70 75 80	
AAA	AAC CGC GTT TCC TTG ACG GTG GAC GGC TTG GCG CAA ATA CAG	430
Lys	Asn Arg Val Ser Leu Thr Val Asp Gly Leu Ala Gln Ile Gln	
	85 90 95	
TCC	TAC ACC GCG CAG GCG GCA TTG GGC GGG ACG AGG ACG GCG GGC	475
Ser	Tyr Thr Ala Gln Ala Ala Leu Gly Gly Thr Arg Thr Ala Gly	
	100 105 110	
AGC	AGC GGC GCA ATC AAT GAA ATC GAG TAT GAA AAC GTC AAA GCT	520
Ser	Ser Gly Ala Ile Asn Glu Ile Glu Tyr Glu Asn Val Lys Ala	
	115 120 125	
GTC	GAA ATC AGC AAA GGC TCA AAC TCG GTC GAA CAA GGC AGC GGC	565
Val	Glu Ile Ser Lys Gly Ser Asn Ser Val Glu Gln Gly Ser Gly	
	130 135 140	

- 37 -

GCA TTG GCG GGT TCG GTC GCA TTT CAA ACC AAA ACC GCC GAC GAT Ala Leu Ala Gly Ser Val Ala Phe Gln Tyr Lys Thr Ala Asp Asp 145 150 155	600
GTT ATC GGG GAA GGC AGG CAG TGG GGC ATT CAG AGT AAA ACC GCC Val Ile Gly Glu Gly Arg Gln Trp Gly Ile Gln Ser Lys Thr Ala 160 165 170	645
TAT TCC GGC AAA AAC CGG GGG CTT ACC CAA TCC ATC GCG CTG GCG Tyr Ser Gly Lys Asn Arg Gly Leu Thr Gln Ser Ile Ala Leu Ala 175 180 185	690
GGG CGC ATC GGC GGT GCG GAG GCT TTG CTG ATC CAC ACC GGG CGG Gly Arg Ile Gly Gly Ala Glu Ala Leu Leu Ile His Thr Gly Arg 190 195 200	735
CGC GCG GGG GAA ATC CGC GCA CAC GAA GAT GCC GGA CGC GGC GTT Arg Ala Gly Glu Ile Arg Ala His Glu Asp Ala Gly Arg Gly Val 205 210 215	780
CAG AGC TTT AAC AGG CTG GTG CCG GTT GAA GAC AGC AGC GAA TAC Gln Ser Phe Asn Arg Leu Val Pro Val Glu Asp Ser Ser Glu Tyr 220 225 230	825
GCC TAT TTC ATC GTT GAA GAT GAA TGC GAA GGC AAA AAT TAC GAA Ala Tyr Phe Ile Val Glu Asp Glu Cys Glu Gly Lys Asn Tyr Glu 235 240 245	870
ACG TGT AAA AGC AAA CCG AAA AAA GAT GTT GTC GGC AAA GAC GAA Thr Cys Lys Ser Lys Pro Lys Lys Asp Val Val Gly Lys Asp Glu 250 255 260	915
CGT CAA ACG GTT TCC ACC CGA GAC TAC ACG GGC CCC AAC CGC TTC Arg Gln Thr Val Ser Thr Arg Asp Tyr Thr Gly Pro Asn Arg Phe 265 270 275	960
CTC GCC GAT CCG CTT TCA TAC GAA AGC CGA TCG TGG CTG TTC CGC Leu Ala Asp Pro Leu Ser Tyr Glu Ser Arg Ser Trp Leu Phe Arg 280 285 290	1005
CCG GGT TTT CGT TTT GAA AAC AAA CGG CAC TAC ATC GGC GGC ATA Pro Gly Phe Arg Phe Glu Asn Lys Arg His Tyr Ile Gly Gly Ile 295 300 305	1050
CTC GAA CAC ACG CAA CAA ACT TTC GAC ACG CGC GAT ATG ACG GTT Leu Glu His Thr Gln Gln Thr Phe Asp Thr Arg Asp Met Thr Val 310 315 320	1095
CCG GCA TTC CTG ACC AAG GCG GTT TTT GAT GCA AAT TCA AAA CAG Pro Ala Phe Leu Thr Lys Ala Val Phe Asp Ala Asn Ser Lys Gln 325 330 335	1140
GCG GGT TCT TTG CCC GGC AAC GGC AAA TAC GCG GGC AAC CAC AAA Ala Gly Ser Leu Pro Gly Asn Gly Lys Tyr Ala Gly Asn His Lys 340 345 350	1185
TAC GGC GGA CTG TTT ACC AAC GGC GAA AAC GGT GCG CTG GTG GGC Tyr Gly Gly Leu Phe Thr Asn Gly Glu Asn Gly Ala Leu Val Gly 355 360 365	1230
GCG GAA TAC GGT ACG GGC GTG TTT TAC GAC GAG ACG CAC ACC AAA Ala Glu Tyr Gly Thr Gly Val Phe Tyr Asp Glu Thr His Thr Lys 370 375 380	1275

- 38 -

AGC CGC TAC GGT TTG GAA TAT GTC TAT ACC AAT GCC GAT AAA GAC Ser Arg Tyr Gly Leu Glu Tyr Val Tyr Thr Asn Ala Asp Lys Asp 385 390 395	1320
ACT TGG GCG GAT TAT GCC CGC CTC TCT TAC GAC CCG CAG GGC ATC Thr Trp Ala Asp Tyr Ala Arg Leu Ser Tyr Asp Arg Gln Gly Ile 400 405 410	1365
GGT TTG GAC AAT CAT TTT CAG CAG ACG CAC TGT TCT GCC GAC GGT Gly Leu Asp Asn His Phe Gln Gln Thr His Cys Ser Ala Asp Gly 415 420 425	1410
TCG GAC AAA TAT TGC CGC CCG AGT GCC GAC AAG CCG TTT TCC TAT Ser Asp Lys Tyr Cys Arg Pro Ser Ala Asp Lys Pro Phe Ser Tyr 430 435 440	1455
TAC AAA TCC GAC CGC GTG ATT TAC GGG GAA AGC CAC AGG CTC TTG Tyr Lys Ser Asp Arg Val Ile Tyr Gly Glu Ser His Arg Leu Leu 445 450 455	1500
CAG GCG GCA TTC AAA AAA TCC TTC GAT ACC GCC AAA ATC CGC CAC Gln Ala Ala Phe Lys Lys Ser Phe Asp Thr Ala Lys Ile Arg His 460 465 470	1545
AAC CTG AGC GTG AAT CTC GGG TTT GAC CGC TTT GAC TCT AAT CTC Asn Leu Ser Val Asn Leu Gly Phe Asp Arg Phe Asp Ser Asn Leu 475 480 485	1590
CGC CAT CAG GAT TAT TAT TAT CAA CAT GCC AAC CGC GCC TAT TCG Arg His Gln Asp Tyr Tyr Tyr Gln His Ala Asn Arg Ala Tyr Ser 490 495 500	1635
TCG AAA ACG CCC CCT AAA ACC GCC AAC CCC AAC GGC GAC AAG AGC Ser Lys Thr Pro Pro Lys Thr Ala Asn Pro Asn Gly Asp Lys Ser 505 510 515	1680
AAA CCC TAT TGG GTC AGC ATA GGC GGG GGA AAT GTG GTT ACG GGG Lys Pro Tyr Trp Val Ser Ile Gly Gly Gly Asn Val Val Thr Gly 520 525 530	1725
CAA ATC TGC CTC TTT GGC AAC AAT ACT TAT ACG GAC TGC ACG CCG Gln Ile Cys Leu Phe Gly Asn Asn Thr Tyr Thr Asp Cys Thr Pro 535 540 545	1770
CGC AGC ATC AAC GGC AAA AGC TAT TAC GCG GCA GTT CCG GAC AAT Arg Ser Ile Asn Gly Lys Ser Tyr Tyr Ala Ala Val Arg Asp Asn 550 555 560	1815
GTC CGT TTG GGC AGG TGG GCG GAT GTC GGC GCG GGG TTG CGC TAC Val Arg Leu Gly Arg Trp Ala Asp Val Gly Ala Gly Leu Arg Tyr 565 570 575	1860
GAC TAC CGC AGC ACG CAT TCG GAC GAC GGC AGC GTT TCC ACC GGC Asp Tyr Arg Ser Thr His Ser Asp Asp Gly Ser Val Ser Thr Gly 580 585 590	1905
ACG CAC CGC ACC CTG TCC TGG AAC GCC GGC ATC GTC CTC AAA CCT Thr His Arg Thr Leu Ser Trp Asn Ala Gly Ile Val Leu Lys Pro 595 600 605	1950
GCC GAC TGG CTG GAT TTG ACT TAC CGC ACT TCA ACC GGC TTC CGC Ala Asp Trp Leu Asp Leu Thr Tyr Arg Thr Ser Thr Gly Phe Arg 610 615 620	1995

CTG CCC TCG TTT GCG GAA ATG TAC GGC TGG CGG TCG GGT GTT CAA Leu Pro Ser Phe Ala Glu Met Tyr Gly Trp Arg Ser Gly Val Gln 625 630 635	2040
AGC AAG GCG GTC AAA ATC GAT CCG GAA AAA TCG TTC AAC AAA GAA Ser Lys Ala Val Lys Ile Asp Pro Glu Lys Ser Phe Asn Lys Glu 640 645 650	2095
GCC GGC ATC GTG TTT AAA GGC GAT TTC GGC AAC TTG GAG GCA AGT Ala Gly Ile Val Phe Lys Gly Asp Phe Gly Asn Leu Glu Ala Ser 655 660 665	2130
TGG TTC AAC AAT GCC TAC CGC GAT TTG ATT GTC CGG GGT TAT GAA Trp Phe Asn Asn Ala Tyr Arg Asp Leu Ile Val Arg Gly Tyr Glu 670 675 680	2175
GCG CAA ATT AAA AAC GGC AAA GAA GAA GCC AAA GGC GAC CCG GCT Ala Gln Ile Lys Asn Gly Lys Glu Glu Ala Lys Gly Asp Pro Ala 685 690 695	2220
TAC CTC AAT GCC CAA AGC GCG CGG ATT ACC GGC ATC AAT ATT TTG Tyr Leu Asn Ala Gln Ser Ala Arg Ile Thr Gly Ile Asn Ile Leu 700 705 710	2265
GGC AAA ATC GAT TGG AAC GGC GTA TGG GAT AAA TTG CCC GAA GGT Gly Lys Ile Asp Trp Asn Gly Val Trp Asp Lys Leu Pro Glu Gly 715 720 725	2310
TGG TAT TCT ACA TTT GCC TAT AAT CGT GTC CAT GTC CGC GAC ATC Trp Tyr Ser Thr Phe Ala Tyr Asn Arg Val His Val Arg Asp Ile 730 735 740	2355
AAA AAA CGC GCA GAC CGC ACC GAT ATT CAA TCA CAC CTG TTT GAT Lys Lys Arg Ala Asp Arg Thr Asp Ile Gln Ser His Leu Phe Asp 745 750 755	2400
GCC ATC CAA CCC TCG CGC TAT GTC GTC GGC TTG GGC TAT GAC CAA Ala Ile Gln Pro Ser Arg Tyr Val Val Gly Leu Gly Tyr Asp Gln 760 765 770	2445
CCG GAA GGC AAA TGG GGT GTG AAC GGT ATG CTG ACT TAT TCC AAA Pro Glu Gly Lys Trp Gly Val Asn Gly Met Leu Thr Tyr Ser Lys 775 780 785	2490
GCC AAG GAA ATC ACA GAG TTG TTG GGC AGC CGG GCT TTG CTC AAC Ala Lys Glu Ile Thr Glu Leu Leu Gly Ser Arg Ala Leu Leu Asn 790 795 800	2535
GGC AAC AGC CGC AAT ACA AAA GCC ACC GCG CGC CGT ACC CGC CCT Gly Asn Ser Arg Asn Thr Lys Ala Thr Ala Arg Arg Thr Arg Pro 805 810 815	2580
TGG TAT ATT GTG GAT GTG TCC GGT TAT TAC ACG ATT AAA AAA CAC Trp Tyr Ile Val Asp Val Ser Gly Tyr Tyr Thr Ile Lys Lys His 820 825 830	2625
TTC ACC CTC CGT GCG GGC GTG TAC AAC CTC CTC AAC TAC CGC TAT Phe Thr Leu Arg Ala Gly Val Tyr Asn Leu Leu Asn Tyr Arg Tyr 835 840 845	2670
GTT ACT TGG GAA AAT GTG CGG CAA ACT GCC GGC GGC GCA GTC AAC Val Thr Trp Glu Asn Val Arg Gln Thr Ala Gly Gly Ala Val Asn 850 855 860	2715

- 40 -

CAA	CAC	AAA	AAT	GTC	GGC	GTT	TAC	AAC	CGA	TAT	GCC	GCC	CCC	GGC	2760
Gln	His	Lys	Asn	Val	Gly	Val	Tyr	Asn	Arg	Tyr	Ala	Ala	Pro	Gly	
			865					870					875		
CGA	AAC	TAC	ACA	TTT	AGC	TTG	GAA	ATG	AAG	TTT	<u>TAA</u>	ACG			2799
Arg	Asn	Tyr	Thr	Phe	Ser	Leu	Glu	Met	Lys	Phe					
			880					885							

SEQ ID NO : 4

Objet : Sequence de l'ADN génomique de la souche de *N. meningitidis* IM2169
codant pour le précurseur de la sous-unité Tbp2 et séquence protéique déduite.
Le peptide signal est indiqué en caractère gras.

ATTTGTTAA AAATAAATAA AATAATAATC CTTATCATTC TTTAATTGAA TTGGGTTTAT	59
<u>ATG</u> AAC AAT CCA TTG GTA AAT CAG GCT GCT ATG GTG CTG CCT GTG	104
Met Asn Asn Pro Leu Val Asn Gln Ala Ala Met Val Leu Pro Val	
-20 -15 -10	
TTT TTG TTG AGT GCC TGT CTG GGC GGC GGC GGC AGT TTC GAT CTT	149
Phe Leu Leu Ser Ala Cys Leu Gly Gly Gly Gly Ser Phe Asp Leu	
-5 -1 1 5 10	
GAT TCT GTC GAT ACC GAA GCC CCG CGT CCC GCG CCA AAG TAT CAA	194
Asp Ser Val Asp Thr Glu Ala Pro Arg Pro Ala Pro Lys Tyr Gln	
15 20 25	
GAT GTT TCT TCC GAA AAA CCG CAA GCC CAA AAA GAC CAA GCC GGA	239
Asp Val Ser Ser Glu Lys Pro Gln Ala Gln Lys Asp Gln Gly Gly	
30 35 40	
TAC GGT TTT GCG ATG AGG TTG AAA CGG AGG AAT TGG TAT CCG GGG	284
Tyr Gly Phe Ala Met Arg Leu Lys Arg Arg Asn Trp Tyr Pro Gly	
45 50 55	
GCA GAA GAA AGC GAG GTT AAA CTG AAC GAG AGT GAT TGG GAG GCG	329
Ala Glu Glu Ser Glu Val Lys Leu Asn Glu Ser Asp Trp Glu Ala	
60 65 70	
ACG GGA TTG CCG ACA AAA CCC AAG GAA CTT CCT AAA CGG CAA AAA	374
Thr Gly Leu Pro Thr Lys Pro Lys Glu Leu Pro Lys Arg Gln Lys	
75 80 85	
TCG GTT ATT GAA AAA GTA GAA ACA GAC GGC GAC AGC GAT ATT TAT	419
Ser Val Ile Glu Lys Val Glu Thr Asp Gly Asp Ser Asp Ile Tyr	
90 95 100	
TCT TCC CCC TAT CTC ACA CCA TCA AAC CAT CAA AAC GGC AGC GCT	464
Ser Ser Pro Tyr Leu Thr Pro Ser Asn His Gln Asn Gly Ser Ala	
105 110 115	
GGC AAC GGT GTA AAT CAA CCT AAA AAT CAG GCA ACA GGT CAC GAA	509
Gly Asn Gly Val Asn Gln Pro Lys Asn Gln Ala Thr Gly His Glu	
120 125 130	
AAT TTC CAA TAT GTT TAT TCC GGT TGG TTT TAT AAA CAT GCA GCG	554
Asn Phe Gln Tyr Val Tyr Ser Gly Trp Phe Tyr Lys His Ala Ala	
135 140 145	
AGT GAA AAA GAT TTC AGT AAC AAA AAA ATT AAG TCA GGC GAC GAT	599
Ser Glu Lys Asp Phe Ser Asn Lys Lys Ile Lys Ser Gly Asp Asp	
150 155 160	

- 42 -

GGT TAT ATC TTC TAT CAC GGT GAA AAA CCT TCC CGA CAA CTT CCT	644
Gly Tyr Ile Phe Tyr His Gly Glu Lys Pro Ser Arg Gln Leu Pro	
165 170 175	
GCT TCT GGA AAA GTT ATC TAC AAA GGT GTG TGG CAT TTT GTA ACC	689
Ala Ser Gly Lys Val Ile Tyr Lys Gly Val Trp His Phe Val Thr	
180 185 190	
GAT ACA AAA AAG GGT CAA GAT TTT CGT GAA ATT ATC CAG CCT TCA	734
Asp Thr Lys Lys Gly Gln Asp Phe Arg Glu Ile Ile Gln Pro Ser	
195 200 205	
AAA AAA CAA GGC GAC AGG TAT AGC GGA TTT TCT GGT GAT GGC AGC	779
Lys Lys Gln Gly Asp Arg Tyr Ser Gly Phe Ser Gly Asp Gly Ser	
210 215 220	
GAA GAA TAT TCC AAC AAA AAC GAA TCC ACG CTG AAA GAT GAT CAC	824
Glu Glu Tyr Ser Asn Lys Asn Glu Ser Thr Leu Lys Asp Asp His	
225 230 235	
GAG GGT TAT GGT TTT ACC TCG AAT TTA GAA GTG GAT TTC GGC AAT	869
Glu Gly Tyr Gly Phe Thr Ser Asn Leu Glu Val Asp Phe Gly Asn	
240 245 250	
AAG AAA TTG ACG GGT AAA TTA ATA CGC AAT AAT GCG AGC CTA AAT	914
Lys Lys Leu Thr Gly Lys Leu Ile Arg Asn Asn Ala Ser Leu Asn	
255 260 265	
AAT AAT ACT AAT AAT GAC AAA CAT ACC ACC CAA TAC TAC AGC CTT	959
Asn Asn Thr Asn Asn Asp Lys His Thr Thr Gln Tyr Tyr Ser Leu	
270 275 280	
GAT GCA CAA ATA ACA GGC AAC CGC TTC AAC GGC ACG GCA ACG GCA	1004
Asp Ala Gln Ile Thr Gly Asn Arg Phe Asn Gly Thr Ala Thr Ala	
285 290 295	
ACT GAC AAA AAA GAG AAT GAA ACC AAA CTA CAT CCC TTT GTT TCC	1049
Thr Asp Lys Lys Glu Asn Glu Thr Lys Leu His Pro Phe Val Ser	
300 305 310	
GAC TCG TCT TCT TTG AGC GGC GGC TTT TTC GGC CCG CAG GGT GAG	1094
Asp Ser Ser Ser Leu Ser Gly Gly Phe Phe Gly Pro Gln Gly Glu	
315 320 325	
GAA TTG GGT TTC CGC TTT TTG AGC GAC GAT CAA AAA GTT GCC GTT	1139
Glu Leu Gly Phe Arg Phe Leu Ser Asp Asp Gln Lys Val Ala Val	
330 335 340	
GTC GGC AGC GCG AAA ACC AAA GAC AAA CTG GAA AAT GGC GCG GCG	1184
Val Gly Ser Ala Lys Thr Lys Asp Lys Leu Glu Asn Gly Ala Ala	
345 350 355	
GCT TCA GGC AGC ACA GGT GCG GCA GCA TCG GGC GGT GCG GCA GGC	1229
Ala Ser Gly Ser Thr Gly Ala Ala Ala Ser Gly Gly Ala Ala Gly	
360 365 370	
ACG TCG TCT GAA AAC AGT AAG CTG ACC ACG GTT TTG GAT GCG GTT	1274
Thr Ser Ser Glu Asn Ser Lys Leu Thr Thr Val Leu Asp Ala Val	
375 380 385	
GAA TTG ACA CTA AAC GAC AAG AAA ATC AAA AAT CTC GAC AAC TTC	1319
Glu Leu Thr Leu Asn Asp Lys Lys Ile Lys Asn Leu Asp Asn Phe	
390 395 400	

AGC AAT GCC GCC CAA CTG GTT GTC GAC GGC ATT ATG ATT CCG CTC Ser Asn Ala Ala Gln Leu Val Val Asp Gly Ile Met Ile Pro Leu 405 410 415	1364
CTG CCC AAG GAT TCC GAA AGC GGG AAC ACT CAG GCA GAT AAA GGT Leu Pro Lys Asp Ser Glu Ser Gly Asn Thr Gln Ala Asp Lys Gly 420 425 430	1409
AAA AAC GGC GGA ACA GAA TTT ACC CGC AAA TTT GAA CAC ACG CCG Lys Asn Gly Gly Thr Glu Phe Thr Arg Lys Phe Glu His Thr Pro 435 440 445	1454
GAA AGT GAT AAA AAA GAC GCC CAA GCA GGT ACG CAG ACG AAT GGG Glu Ser Asp Lys Lys Asp Ala Gln Ala Gly Thr Gln Thr Asn Gly 450 455 460	1499
GCG CAA ACC GCT TCA AAT ACG GCA GGT GAT ACC AAT GGC AAA ACA Ala Gln Thr Ala Ser Asn Thr Ala Gly Asp Thr Asn Gly Lys Thr 465 470 475	1544
AAA ACC TAT GAA GTC GAA GTC TGC TGT TCC AAC CTC AAT TAT CTG Lys Thr Tyr Glu Val Glu Val Cys Cys Ser Asn Leu Asn Tyr Leu 480 485 490	1589
AAA TAC GGA ATG TTG ACG CGC AAA AAC AGC AAG TCC GCG ATG CAG Lys Tyr Gly Met Leu Thr Arg Lys Asn Ser Lys Ser Ala Met Gln 495 500 505	1634
GCA GGA GGA AAC AGT AGT CAA GCT GAT GCT AAA ACG GAA CAA GTT Ala Gly Gly Asn Ser Ser Gln Ala Asp Ala Lys Thr Glu Gln Val 510 515 520	1679
GAA CAA AGT ATG TTC CTC CAA GGC GAG CGT ACC GAT GAA AAA GAG Glu Gln Ser Met Phe Leu Gln Gly Glu Arg Thr Asp Glu Lys Glu 525 530 535	1724
ATT CCA ACC GAC CAA AAC GTC GTT TAT CGG GGG TCT TGG TAC GGG Ile Pro Thr Asp Gln Asn Val Val Tyr Arg Gly Ser Trp Tyr Gly 540 545 550	1769
CAT ATT GCC AAC GGC ACA AGC TGG AGC GGC AAT GCT TCT GAT AAA His Ile Ala Asn Gly Thr Ser Trp Ser Gly Asn Ala Ser Asp Lys 555 560 565	1814
GAG GGC GGC AAC AGG GCG GAA TTT ACT GTG AAT TTT GCC GAT AAA Glu Gly Gly Asn Arg Ala Glu Phe Thr Val Asn Phe Ala Asp Lys 570 575 580	1859
AAA ATT ACC GGC AAG TTA ACC GCT GAA AAC AGG CAG GCG CAA ACC Lys Ile Thr Gly Lys Leu Thr Ala Glu Asn Arg Gln Ala Gln Thr 585 590 595	1904
TTT ACC ATT GAG GGA ATG ATT CAG GGC AAC GGC TTT GAA GGT ACG Phe Thr Ile Glu Gly Met Ile Gln Gly Asn Gly Phe Glu Gly Thr 600 605 610	1949
GCG AAA ACT GCT GAG TCA GGT TTT GAT CTC GAT CAA AAA AAT ACC Ala Lys Thr Ala Glu Ser Gly Phe Asp Leu Asp Gln Lys Asn Thr 615 620 625	1994
ACC CGC ACG CCT AAG GCA TAT ATC ACA GAT GCC AAG GTA AAG GGC Thr Arg Thr Pro Lys Ala Tyr Ile Thr Asp Ala Lys Val Lys Gly 630 635 640	2039

- 44 -

GGT TTT TAC GGG CCT AAA GCC GAA GAG TTG GGC GGA TGG TTT GCC	2084
Gly Phe Tyr Gly Pro Lys Ala Glu Glu Leu Gly Gly Trp Phe Ala	
645 650 655	
TAT CCG GGC GAT AAA CAA ACG GAA AAG GCA ACA GCT ACA TCC AGC	2129
Tyr Pro Gly Asp Lys Gln Thr Glu Lys Ala Thr Ala Thr Ser Ser	
660 665 670	
GAT GGA AAT TCA GCA AGC AGC GCG ACC GTG GTA TTC GGT GCG AAA	2174
Asp Gly Asn Ser Ala Ser Ser Ala Thr Val Val Phe Gly Ala Lys	
675 680 685	
CGC CAA CAG CCT GTG CAA <u>TAA</u> GCACGGTTGC CGAACAAATCA AGAATAAGGC	2225
Arg Gln Gln Pro Val Gln	
690	
TTCAG	2230

Revendications

1. Un fragment d'ADN isolé codant pour un peptide, un polypeptide ou une protéine capables d'être reconnus par un antisérum anti-récepteur de la transferrine de la souche de *N. meningitidis* IM2394 ou IM2169.
2. Un fragment d'ADN selon la revendication 1, qui comprend une séquence nucléotidique codant pour une séquence d'acides aminés homologue à celle telle que montrée :
 - dans le SEQ ID NO : 1, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579 ;
 - dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 884 ;
 - dans le SEQ ID NO : 3, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887 ;
ou
 - dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691.
3. Un fragment d'ADN selon la revendication 2, qui comprend une séquence nucléotidique codant pour une séquence d'acides aminés telle que montrée :
 - dans le SEQ ID NO : 1, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579 ;
 - dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 884 ;
 - dans le SEQ ID NO : 3, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887 ;
ou

- dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691.
4. Un fragment d'ADN selon la revendication 2, qui a une séquence nucléotidique codant pour une protéine ayant une séquence d'acides aminés homologue à celle telle que montrée :
- dans le SEQ ID NO : 1, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579 ;
 - dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 884 ;
 - dans le SEQ ID NO : 3, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887 ;
ou
 - dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691.
5. Un fragment d'ADN selon la revendication 4, qui a une séquence nucléotidique codant pour :
- i) la sous-unité Tbp1 de la souche IM2394 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 2, commençant avec le résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 884 ;
 - ii) la sous-unité Tbp2 de la souche IM2394 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 1, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 579 ;
 - iii) la sous-unité Tbp1 de la souche IM2169 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 3, commençant avec le

résidu acide glutamique en position 1 et finissant avec le résidu phénylalanine en position 887 ; ou

- iv) la sous-unité Tbp2 de la souche IM2169 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 4, commençant avec le résidu cystéine en position 1 et finissant avec le résidu glutamine en position 691.
6. Un fragment d'ADN selon la revendication 2, qui a une séquence nucléotidique codant pour un précurseur ayant une séquence d'acides aminés homologue à celle telle que montrée :
- dans le SEQ ID NO : 1, commençant avec le résidu méthionine en position -20 et finissant avec le résidu glutamine en position 579 ;
 - dans le SEQ ID NO : 2, commençant avec le résidu méthionine en position -24 et finissant avec le résidu phénylalanine en position 884 ;
 - dans le SEQ ID NO : 3, commençant avec le résidu méthionine en position -24 et finissant avec le résidu phénylalanine en position 887 ;
ou
 - dans le SEQ ID NO : 4, commençant avec le résidu méthionine en position -20 et finissant avec le résidu glutamine en position 691.
7. Un fragment d'ADN selon l'une des revendications 1 et 2, qui a une séquence nucléotidique codant pour :
- i) Le précurseur de la sous-unité Tbp1 de la souche IM2394 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 2, commençant avec le résidu méthionine en position -24 et finissant avec le résidu phénylalanine en position 884 ;
 - ii) Le précurseur de la sous-unité Tbp2 de la souche IM2394 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 1, commençant avec le résidu méthionine en position -20 et finissant

avec le résidu glutamine en position 579 ;

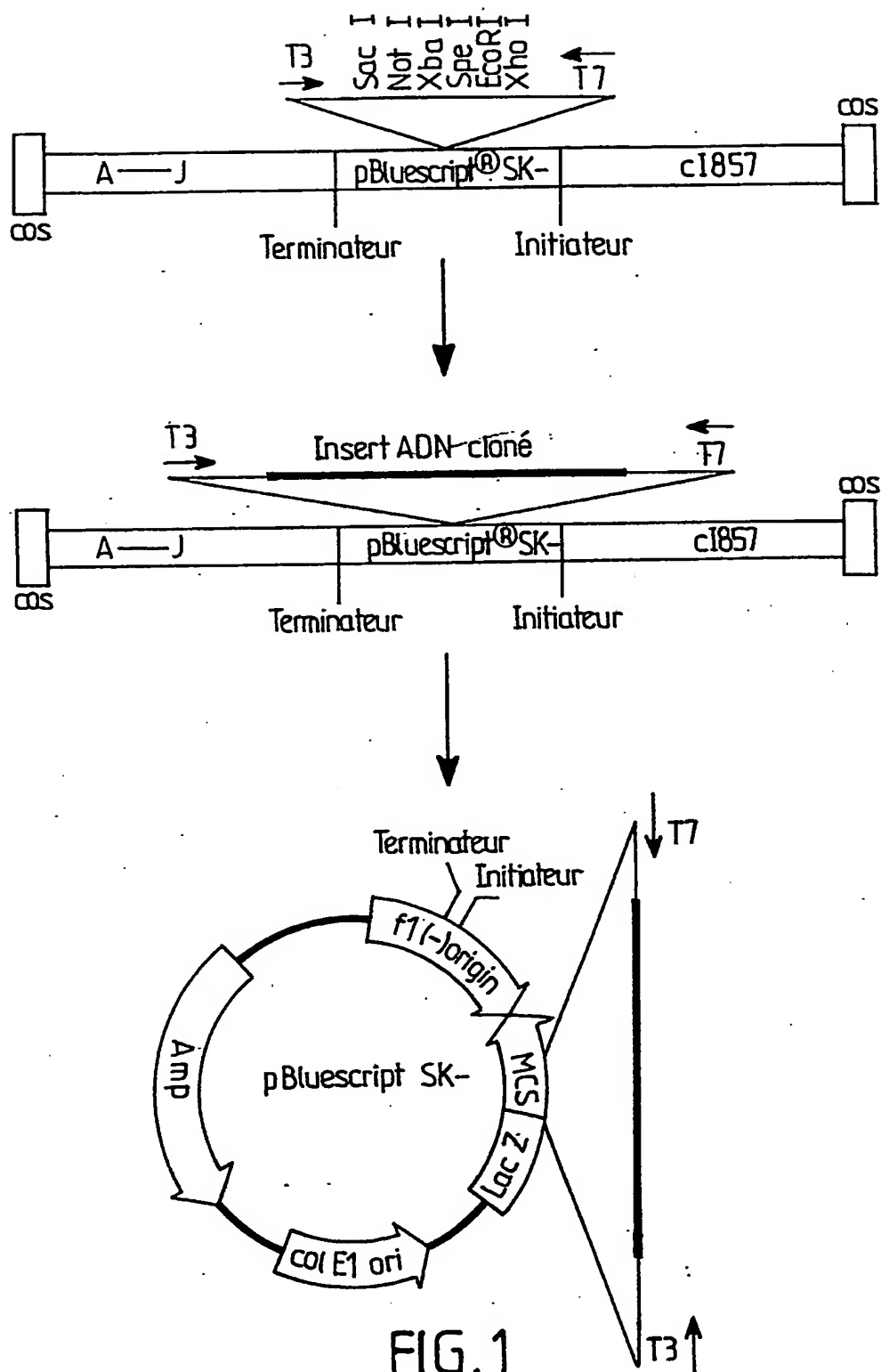
- iii) Le précurseur de la sous-unité Tbp1 de la souche IM2169 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 3, commençant avec le résidu méthionine en position -24 et finissant avec le résidu phénylalanine en position 887 ; ou
 - iv) Le précurseur de la sous-unité Tbp2 de la souche IM2169 dont la séquence en acides aminés est montrée dans le SEQ ID NO : 4, commençant avec le résidu méthionine en position -20 et finissant avec le résidu glutamine en position 691.
8. Une cassette d'expression destinée à la production d'une protéine capable d'être reconnue par un antisérum anti-récepteur de la transferrine de la souche de *N. meningitidis* IM2394 ou IM2169, qui comprend un fragment d'ADN selon l'une des revendications 1 à 7, placé sous le contrôle des éléments nécessaires à son expression.
 9. Une cellule-hôte transformée par une cassette d'expression selon la revendication 8.
 10. Un procédé de production d'une protéine capable d'être reconnue par un antisérum anti-récepteur de la transferrine de la souche de *N. meningitidis* IM2394 ou IM2169, qui comprend l'acte de cultiver une cellule-hôte selon la revendication 9.
 11. Un bloc d'ADN isolé codant pour un peptide signal ayant une séquence d'acides aminés homologue à celle telle que montrée dans :
 - le SEQ ID NO : 2, commençant avec le résidu méthionine en position - 24 et finissant avec le résidu en position - 1.
 - le SEQ ID NO : 3, commençant avec le résidu méthionine en position - 24 et finissant avec le résidu alanine en position - 1 ; et
 - le SEQ ID NO : 4, commençant avec le résidu méthionine en

position - 20 et finissant avec le résidu alanine en position - 1.

12. Un bloc d'ADN isolé codant pour un peptide signal ayant une séquence d'acides aminés telle que montrée dans :

- le SEQ ID NO : 2, commençant avec le résidu méthionine en position - 24 et finissant avec le résidu en position - 1 ;
- le SEQ ID NO : 3, commençant avec le résidu méthionine en position - 24 et finissant avec le résidu alanine en position - 1 ; et
- le SEQ ID NO : 4, commençant avec le résidu méthionine en position - 20 et finissant avec le résidu alanine en position - 1.

1/10



2/10

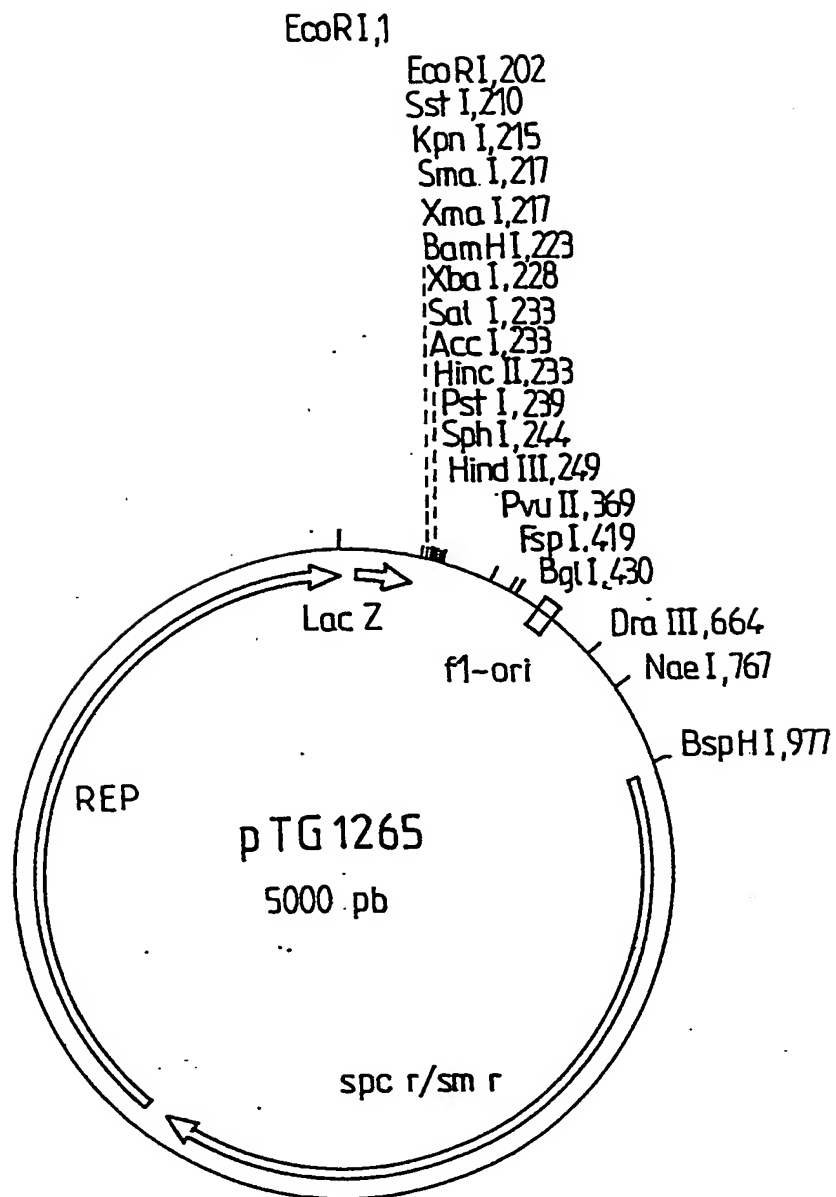


FIG. 2

3/10

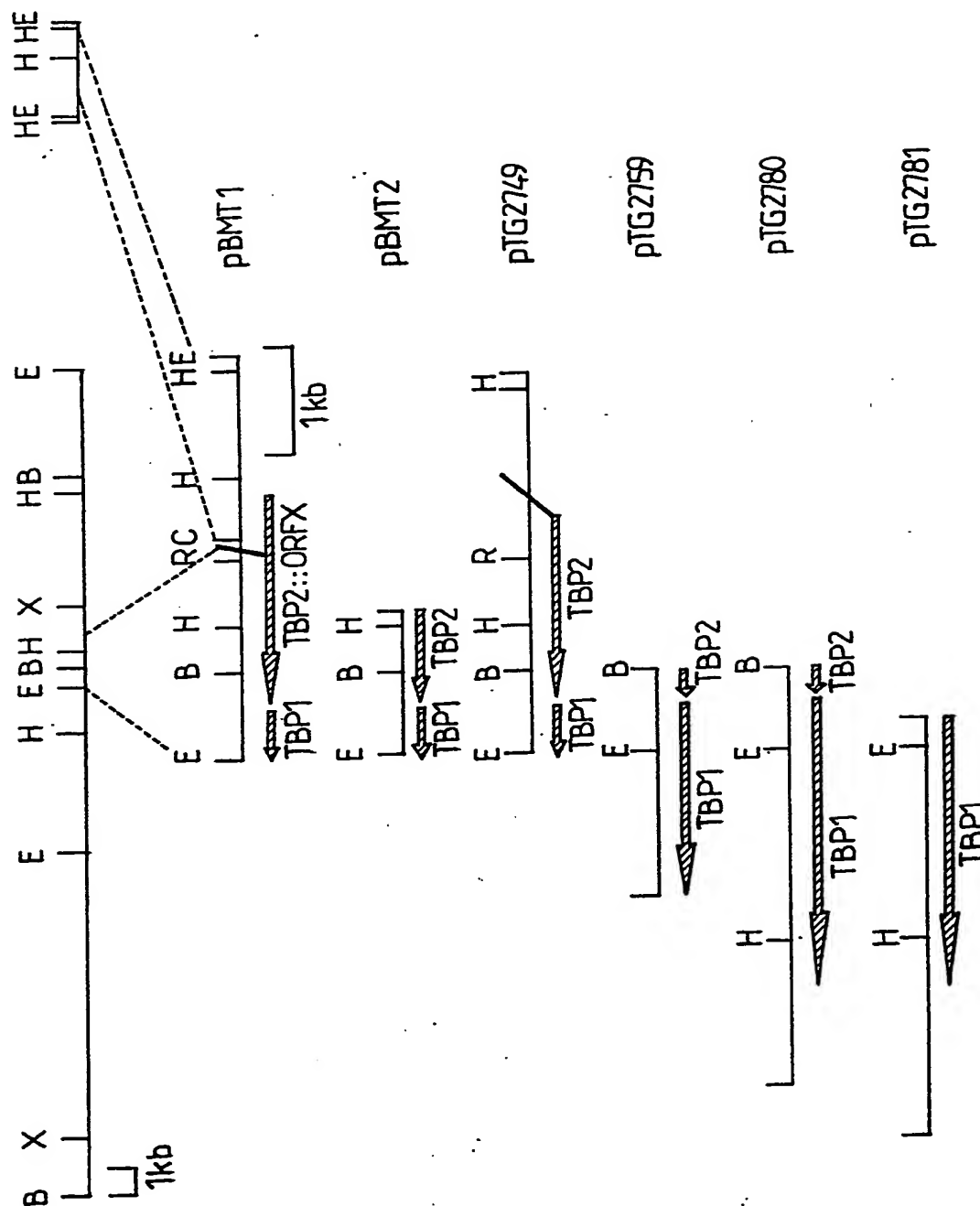


FIG. 3

4/10

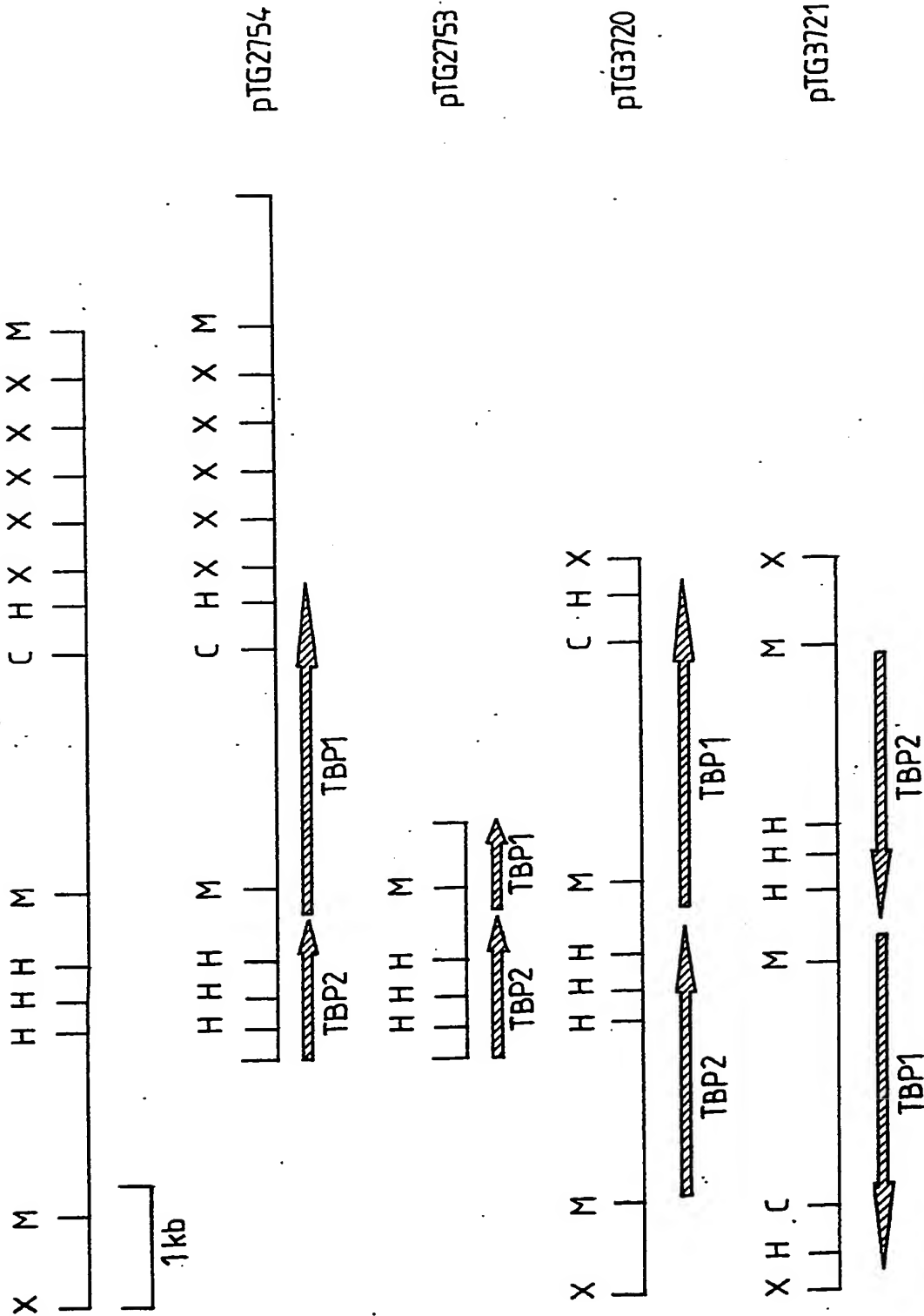


FIG. 4

5/10

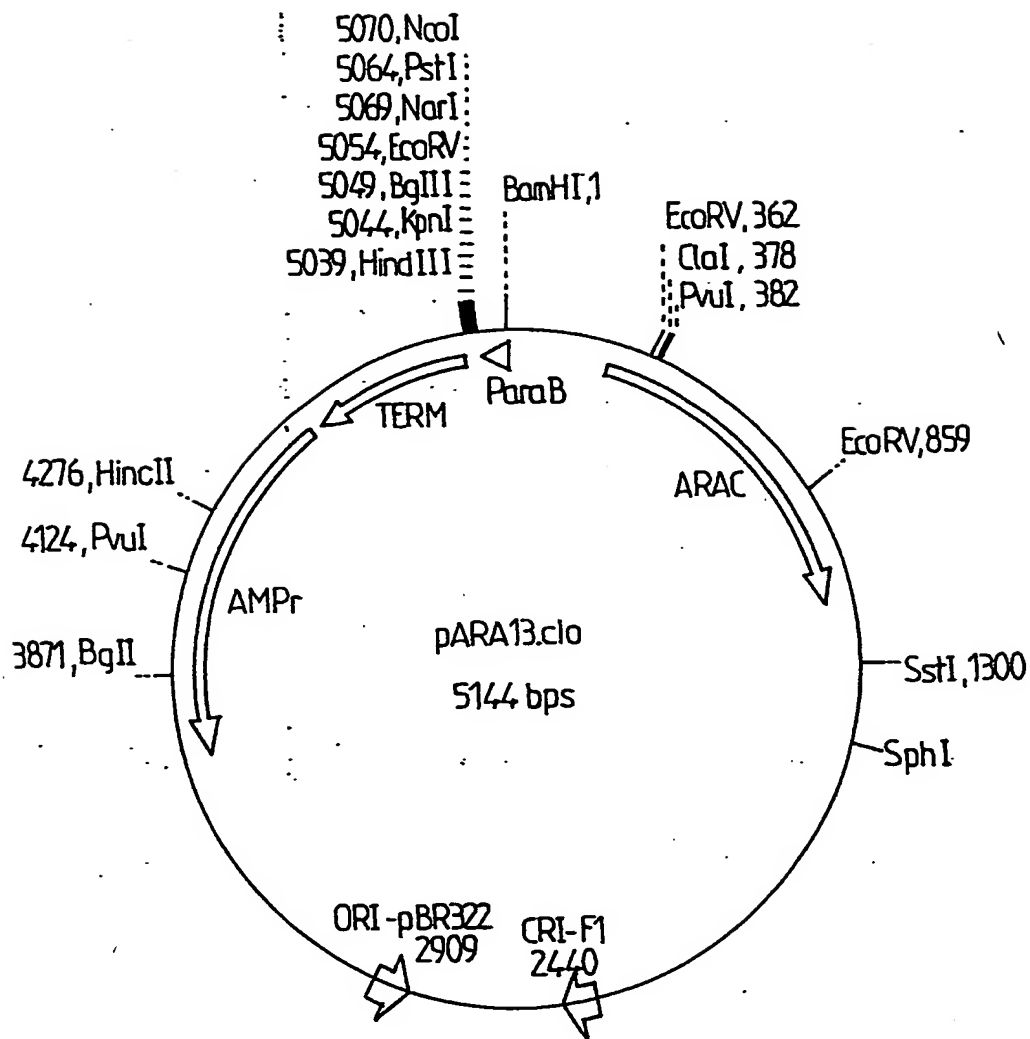


FIG. 5

6/10

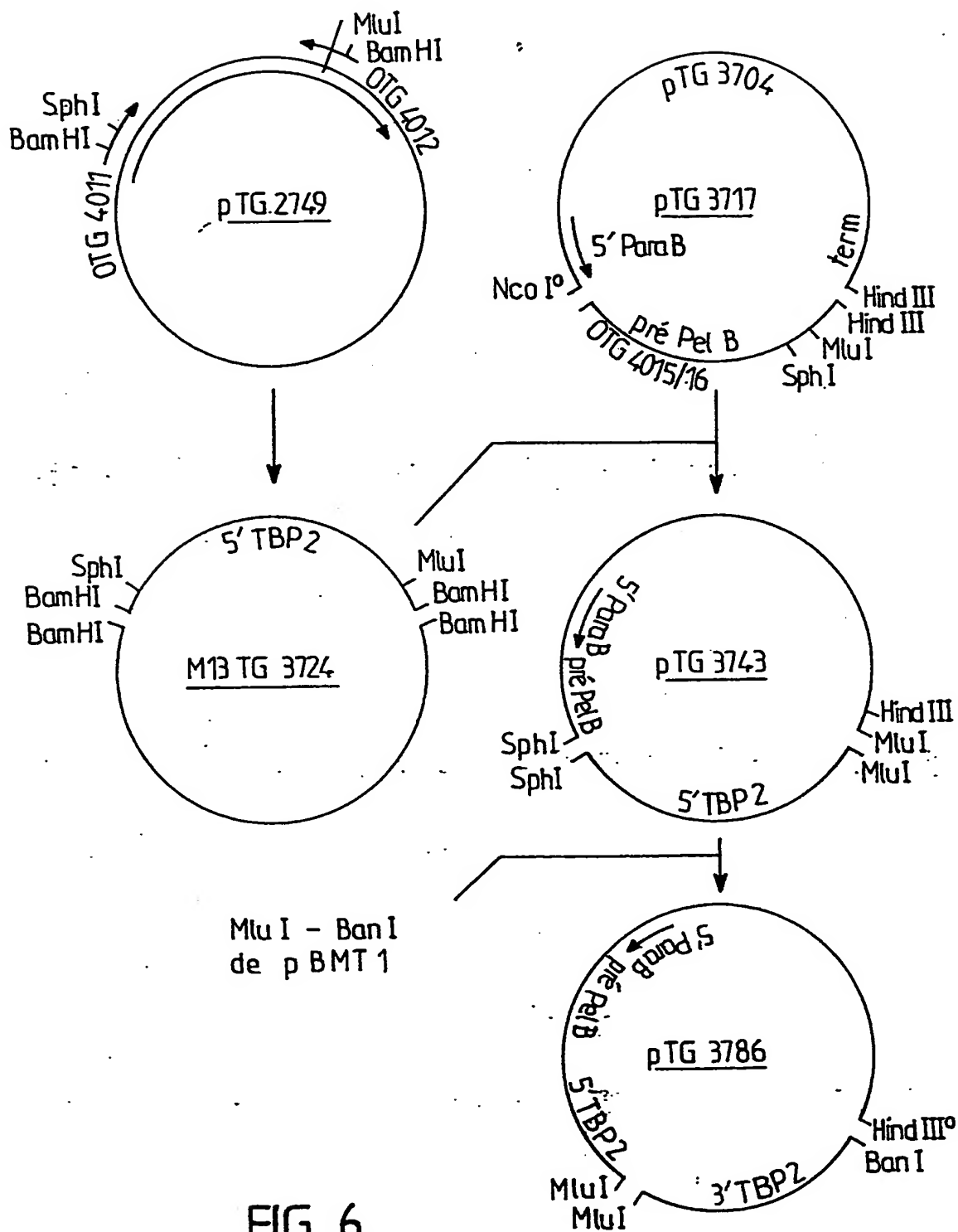


FIG. 6

7/10

Tbp1-2394	MQQQHLEFRLNLCLSLMTALPVYAENVQAEQAEQXQLDTIQVKAKKQKTRRDNEVTGLGK
Tbp1-2169	MQQQHLEFRLNLCLSLMTALPAYAENVQAGQAEQXQLDTIQVKAKKQKTRRDNEVTGLGK
Tbp1-2394	LVKSSDTLSKEQVLNIRDLTRYDPGLIAVVEQGRGASSGYSIRGMDKNRVSLTVDGVSQIQ
Tbp1-2169	LVKTADTLSKEQVLDIRDLTRYDPGLIAVVEQGRGASSGYSIRGMDKNRVSLTVDGLAQIQ
Tbp1-2394	SYTAQAALGGTRTAGSSGAINIEIYENVKAVEISKGSNSSEYGNAGLAGSVAFQTKTAAD
Tbp1-2169	SYTAQAALGGTRTAGSSGAINIEIYENVKAVEISKGSNSVEQSGALAGSVAFQTKTADO
Tbp1-2394	IIGEGKQWGIQSKTAYSGKDHALTQSLALAGRSGGAEALLIYTKRRGREIHAKDAGKGV
Tbp1-2169	VIGEGROWGIQSKTAYSGKNRGLTQSLALAGRIGGAEALLIHTGRRAGEIRAHEDAGRGV
Tbp1-2394	QSFNRLVLDDEKKEGGSQYRYFIVEEECH-NGYAAACKNKLKEDASVKDERKTVSTQOYTG
Tbp1-2169	QSFNRLVPVED-----SSEYAYFIVEDECEGKNYETCKSKPKDVGKDERQTVSTRDYTG
Tbp1-2394	SNRLLANPLEYGSQSWLFRPGWHLDN-RHYVGAVLERTQQTFTDTRDMTPPAYFTSEDYVP
Tbp1-2169	PNRFLADPLSYESRSWLFPGFRFENKRHYIGGILEZHTQQTFTDTRDMTPPAFLTKAVFDA
Tbp1-2394	-----GSLKGLGKYSGONKAERLFVQEGGSTLQIGYGTGVFYDERHTKNRYGVEYVYHN
Tbp1-2169	NSKQAGSLPGNGKYAGNHXYGGLFTNGENGALVGAEYGTGVFYDETHKSRYLEYVYTN
Tbp1-2394	ADKDTWADYARLSYDRQGIIDLNRLLQQTHCSHDGSDKNCRPOGNKPYSFYKSDRMTYEES
Tbp1-2169	ADKDTWADYARLSYDRQGIIDLNRLLQQTHCSADGSDKYCRPSADKPFYKYSDRVITYGES
Tbp1-2394	RNLFAQVEFKAFDTAKIRHNLSINLG'YDRFKSQLSHSDYYLQNAVQAYDLITPKKPPFPN
Tbp1-2169	HRLLQAAFKKSFDTAKIRHNLSVNLG'YDRFKSQLSHSDYYLQNAVQAYDLITPKKPPFPN
Tbp1-2394	GSKDNPYRVSIGKTTVNTSPICRFGNNTYTDCTPRNIGGNGYAAVQONVRLGRWADVGA
Tbp1-2169	GDKSKPYWVSIGGGNVVTCQICLFGNNTYTDCTPRNIGGNGYAAVQONVRLGRWADVGA

FIG.7

8/10

Tbp1-2394	GIRYDYRSTHSEDKSVSTGTHRNLSWNAGVVLKPF TWMDLTYRASTGFRLPSFAEMYGWR
Tbp1-2169	GLRYDYRSTHSDG SVSTGTHRTL SWNAGIVLKPADWL DLT YRTSTGFRLPSFAEMYGWR
Tbp1-2394	AGESLKTLDLKPEKSFNREAGIVFKGDFGNLEASYFNNAYRDLIAFGYETRTQNGQTSAS
Tbp1-2169	SGVQSKAVKIDPEKSFNKEAGIVFKGDFGNLEASWFNNAYRDLIVRGYEAQIKNGKEEAK
Tbp1-2394	GDPGYRNAQNARIAGINILGKIDWHGVWGGLPDGLYSTLAYNRIKVKDAOIRADRTFVTS
Tbp1-2169	GDPAYLNAQSARITGINILGKIDWNGVWDKLPEGWYSTFAYNRVHVVDIKKRADRTDIQS
Tbp1-2394	YLFDAVQPSRYVLGLGYDHPDGIWGINIMFTYSKAKSVDELLGSQALLNGNANAKKAASR
Tbp1-2169	HLFDAIQPSRYVVGGLGYDQPEGWGVNGMLTYSKAKEITELLGSRALINGNSRNTKATAR
Tbp1-2394	RTRPWYVTOVSGYYNIKKHLTLRAGVYNLLNYRYVTWENVVRQTAGGAVNQHKNVGVYNRY
Tbp1-2169	RTRPWYIVDVSGYYTIKKHFTLRAGVYNLLNYRYVTWENVVRQTAGGAVNQHKNVGVYNRY
Tbp1-2394	AAPGRNYTFSLEMKF
Tbp1-2169	AAPGRNYTFSLEMKF

* = acide aminé identique
 . = changement conservatif

FIG: 7 (suite)

9/10

Tbp2-2394	CLGGGGSFDLDSVETVCCMHKPKYEDKESQ-PESQQQVSENSGAAYGFAVKLPRRNAHF
Tbp2-2169	CLGGGGSFDLDSVDT-EAPRPAPKYQDVSSSEKQPAQKD----QG-GYGFAMRLKRRN--W
Tbp2-2394	NPXYKXKXKPLGSMDWKXLO-RGEPNSFSERDE--LEKKR--G-----SSE-LIESKWED
Tbp2-2169	YPGAEESEVKLNESEDWEATGLPTKPKELPKRQKSVIEKVEITDGDSDIYSSPYLTPSNHQ
Tbp2-2394	G-----Q--SRVVGYNFTYVRSQVYLNK--NNIDIKNNIVLFGPDGYLYYKGEPSK
Tbp2-2169	GSAGNGVNPKNQATGHENFQYVYSGWFYKHAASEKDFSNKKIKSGDDGYIFYHGEKPSR
Tbp2-2394	ELP-SEKITXKGTWQVYTDAMEKQRF-EG--GSAAGGDKSGALSALLEGVLRNQAEAS--
Tbp2-2169	QLPASGKVIYKGVWHEVTDTKKQDFREIIQPSKKQGDRIYSGFSGDGSEYSNKNNESTLK
Tbp2-2394	SGHTDFGMTSEFEVDPSDKTIKGTLYRNNRITQNNSENKQIKTTRYTIQATLHGNRFKXG
Tbp2-2169	DDHEGYGFTSNLEVDGCKKLTGKLRNNASLNNNTNNDKHTTQYYSLDAQITGNRENGT
Tbp2-2394	ALAADKGATNGS--HPFISDSLSLEGGFYGPKEZELAGKFLSNDNKVAAVFGAKQKDKDG
Tbp2-2169	ATATDKKENETKLHPFVSOSSSLSGGFFGPQGEELGFRFLSDQKVAVVGSAKTDKKLEN
Tbp2-2394	EN-----AAGPATE-----TVIDAYRITGEEFKKEQIDSPGDVKKLLVDGVE
Tbp2-2169	GAAASGSTGAAASGGAAGTSSSENSKLTIVLDAVELTLNDKKIKNLDNFSNAAQLVVDGIM
Tbp2-2394	LSLLPSEGKAA-----FQHEIE-----QNGVKAT-----
Tbp2-2169	EPLLPKDSSESGNTQADKKGNGGTEFTTRKFENTPESDKKDAQAGTQTNGAQTASNTAGDTN
Tbp2-2394	-----VCCSNLDVMSFGKLSKEN-K-----DDMFLOQGVRT
Tbp2-2169	GKTKTYEVEVCCSNLNLKYGNLTKNSKSAHQAGGNSSQADAKTEQVEQSMFLOQERTD
Tbp2-2394	VSDVAARTEANAKYRGTYGYIANGTSWSGEASNQEGGNRAEFDVDFSTKKISGTLTAKD
Tbp2-2169	EKEIP--TDQNVVYRGSWYGHIANGTWSGNASDKEGGNRAEFTVNFADKKITGKLIEN
Tbp2-2394	RTSPAFTITAMIKDNGFSGVAKTCENGFAIDPONTGNSHYTHI-EATVSGGFYGNKAIEM
Tbp2-2169	RQAQITFTIEGMIQNGFEGTAKTAESGFDLDQKNTTRTPKAYITDAKVKGFFYGPKAEEI
Tbp2-2394	GGSFSEFPGNAPEGKQE-----KASVVFAGAKROQLVQ
Tbp2-2169	GGWFAYPGDKQTEKATATSSDGNSASSATVVFAGAKROQPVQ

* = acide aminé identique
 . = changement conservatif

FIG. 8

10/10

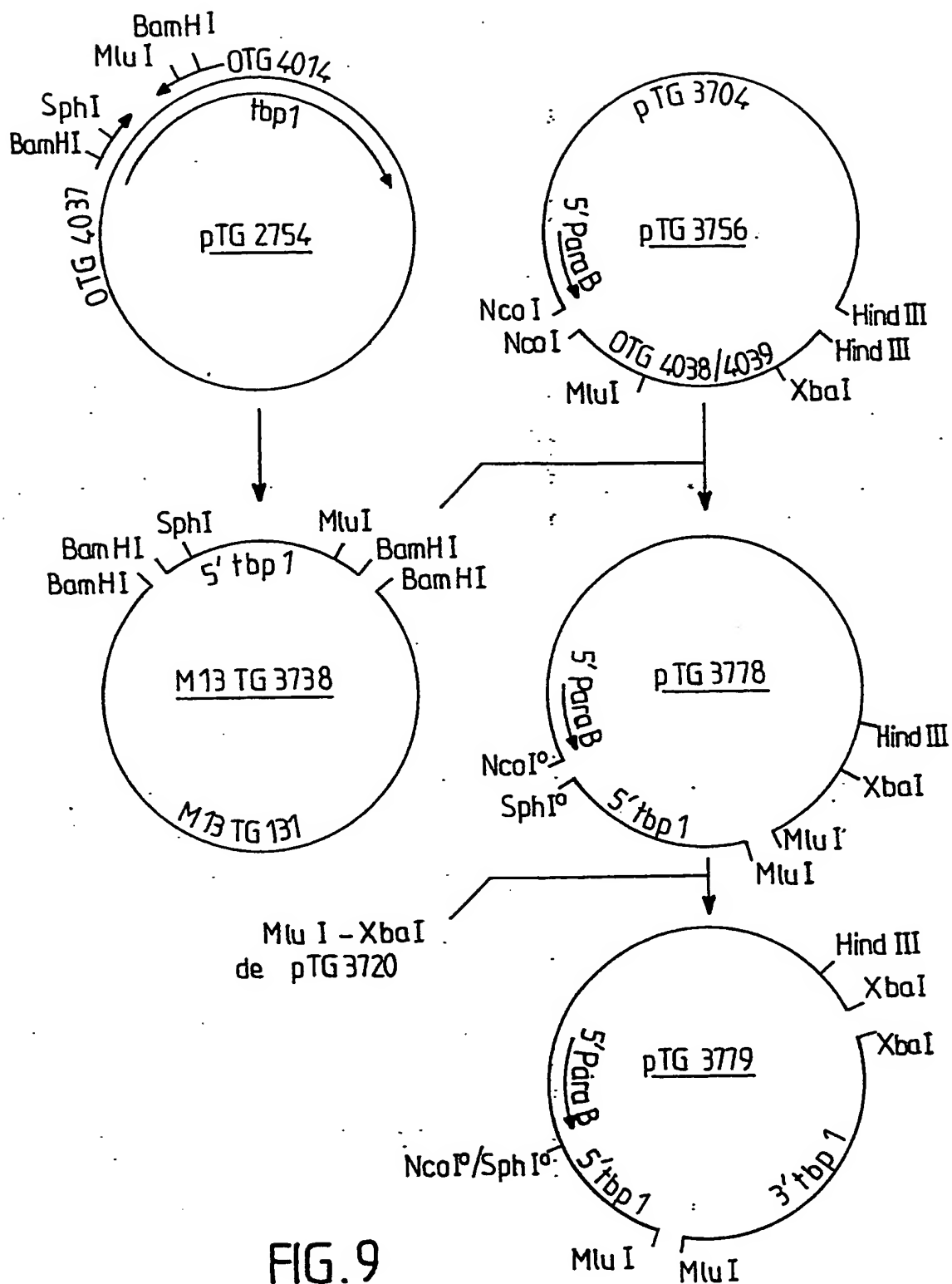


FIG. 9

INSTITUT NATIONAL
de la
PROPRIÉTÉ INDUSTRIELLE

RAPPORT DE RECHERCHE
établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FR 9207493
FA 474262

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
X	INFECTION AND IMMUNITY vol. 60, no. 6, Juin 1992, pages 2391 - 2396 Stevenson P; Williams P; Griffiths E; 'Common antigenic domains in transferrin -binding protein 2 of Neisseria -Meningitidis Neisseria -Gonorrhoeae and Haemophilus-Influenzae Type B.' * le document en entier *	1
D,X	WO-A-9 012 591 (UNIVERSITY TECHNOLOGIES INTERNATIONAL, INC.; US) 1 Novembre 1990 * le document en entier *	1
X	WO-A-9 203 467 (THE UNIVERSITY OF NORTH CAROLINA, US) 5 Mars 1992 * revendications 1-17; figure 2A *	1,11,12
		DOMAINES TECHNIQUES RECHERCHES (Int. Cl.5)
		C12N C07K
Date d'achèvement de la recherche 28 JANVIER 1993		Examinateur S.A. NAUCHE
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>		

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **INES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.